

STIC-Biotech/ChemLib

197582

Fr m: Ramirez, Delia
S nt: Thursday, August 03, 2006 4:56 PM
To: STIC-Biotech/ChemLib
Subject: 10/612779

Hi,

I would like to request the following search: SEQ ID NO:2 in the protein databases.

Please provide a printout of this search.

Thank you very much,

Delia M. Ramirez, Ph.D.
Patent Examiner
Recombinant Enzymes-Art Unit 1652
USPTO
400 Dulany Street, Remsen Bldg., 2D74, Mail room 2C70
Alexandria, VA 22314
(571) 272-0938
delia.ramirez@uspto.gov

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 7, 2006, 09:20:34 ; Search time 92 Seconds
(without alignments)
3026.574 Million cell updates/sec

Title: US-10-612-779-2

Perfect score: 3086

Sequence: 1 MCGIVGAIQRDVAEILLEG.....LRGTVDQPNRLAKSVTVE 609

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

- A_Geneseq_8:*
- 1: Geneseq1980s:*
 - 2: Geneseq1990s:*
 - 3: Geneseq2000s:*
 - 4: Geneseq2001s:*
 - 5: Geneseq2002s:*
 - 6: Geneseq2003as:*
 - 7: Geneseq2003bs:*
 - 8: Geneseq2004s:*
 - 9: Geneseq2005s:*
 - 10: Geneseq2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 3086 | 100.0 | 609 | 3 | AAV58822 E. coli g |
| 2 | 3086 | 100.0 | 609 | 4 | AAU34806 E. coli c |
| 3 | 3086 | 100.0 | 609 | 6 | ABU28829 Protein e |
| 4 | 3086 | 100.0 | 609 | 8 | ADI38857 Glucosami |
| 5 | 3086 | 100.0 | 609 | 8 | ADS45181 Bacterial |
| 6 | 3086 | 100.0 | 609 | 8 | ADU00590 Amino aci |
| 7 | 3081 | 99.8 | 608 | 9 | ADW23840 Novel hum |
| 8 | 3080 | 99.8 | 609 | 3 | AAV58827 E. coli g |
| 9 | 3080 | 99.8 | 609 | 3 | AAV58826 E. coli g |
| 10 | 3080 | 99.8 | 609 | 8 | ADI38865 Mutant gl |
| 11 | 3080 | 99.8 | 609 | 8 | ADI38867 Mutant gl |
| 12 | 3079 | 99.8 | 609 | 3 | AAV58825 E. coli g |
| 13 | 3079 | 99.8 | 609 | 8 | ADI38863 Mutant gl |
| 14 | 3071 | 99.5 | 609 | 3 | AAV58823 E. coli g |
| 15 | 3071 | 99.5 | 609 | 8 | ADI38859 Mutant gl |
| 16 | 3068 | 99.4 | 609 | 3 | AAV58824 E. coli g |
| 17 | 3068 | 99.4 | 609 | 8 | ADI38861 Mutant gl |
| 18 | 3068 | 99.4 | 609 | 8 | ADU00592 Amino aci |
| 19 | 3061 | 99.2 | 609 | 8 | ADI38869 Mutant gl |
| 20 | 3053 | 98.9 | 609 | 4 | AAU38472 Salmonell |
| 21 | 3053 | 98.9 | 609 | 6 | ABU48228 Protein e |
| 22 | 2950 | 95.6 | 609 | 6 | ABU28355 Protein e |
| 23 | 2876 | 93.2 | 609 | 4 | AAU36165 Klebsiell |

| | | | | | |
|----|--------|------|-----|----|--------------------|
| 24 | 2876 | 93.2 | 651 | 7 | ABO64174 Klebsiell |
| 25 | 2871 | 93.0 | 609 | 6 | ABU31411 Protein e |
| 26 | 2718.5 | 88.1 | 608 | 6 | ABU45041 Protein e |
| 27 | 2682 | 86.9 | 609 | 6 | ABU50609 Protein e |
| 28 | 2661 | 86.2 | 609 | 8 | ADS42812 Bacterial |
| 29 | 2619 | 84.9 | 609 | 8 | ADN17785 Bacterial |
| 30 | 2570 | 83.3 | 610 | 6 | ABM70565 Phototrab |
| 31 | 2510.5 | 81.4 | 622 | 7 | ADF07237 Bacterial |
| 32 | 2505.5 | 81.2 | 608 | 6 | ABU41104 Protein e |
| 33 | 2270.5 | 73.6 | 610 | 4 | AAU35425 Haemophil |
| 34 | 2270.5 | 73.6 | 610 | 6 | ABU30239 Protein e |
| 35 | 2270.5 | 73.6 | 610 | 6 | AAE30459 Haemophil |
| 36 | 2270.5 | 73.6 | 610 | 10 | AAE31344 Haemophil |
| 37 | 2264.5 | 73.4 | 610 | 6 | ABU49139 Protein e |
| 38 | 2202.5 | 71.4 | 610 | 6 | ABU39464 Protein e |
| 39 | 2007 | 65.0 | 611 | 6 | ABU39730 Protein e |
| 40 | 2002.5 | 64.9 | 610 | 8 | ADS44814 Bacterial |
| 41 | 1995 | 64.6 | 611 | 4 | AAU36518 Pseudomon |
| 42 | 1995 | 64.6 | 611 | 6 | ABU38924 Protein e |
| 43 | 1995 | 64.6 | 616 | 7 | ABO70418 Pseudomon |
| 44 | 1987 | 64.4 | 611 | 6 | ABU41714 Protein e |
| 45 | 1806.5 | 58.5 | 612 | 6 | ABU17019 Protein e |

ALIGNMENTS

RESULT 1
AAV58822
ID AAY58822 standard; protein; 609 AA.
XX
AC AAY58822;
XX
DT 08-MAY-2000 (first entry)
XX
DE E. coli glucosamine-6-phosphate synthase.
XX
KW Glucosamine-6-phosphate synthase; GlcN6P; glucosamine;
KW metabolic engineering; plasmid pKLN23-28.
XX
OS Escherichia coli.
XX
PN WO200004182-A1.
XX
PD 27-JAN-2000.
XX
PF 15-JUL-1999; 99WO-US015976.
XX
PR 15-JUL-1998; 98US-00115475.
XX
(DCVB-) DCV INC DBA BIO-TECH RESOURCES.
XX
PI Berry A, Burlingame RP, Millis JR;
XX WPI; 2000-182441/16.
DR N-PSDB; AAZ58249.
XX
PT Fermentation of E. coli having an altered amino acid sugar metabolic
PT pathway to produce glucosamine, especially using novel recombinant
variant glucosamine-6-phosphate synthases.
XX
PS Claim 15; Page 111-113; 150pp; English.
XX
CC The present sequence is that of wild-type glucosamine-6-phosphate
CC synthase (GlcN6P synthase) of Escherichia coli strain W3110. The sequence
CC is predicted from the isolated glms gene (see AAZ58249) in plasmid pKLN23
CC -28. Recombinant nucleic acids encoding GlcN6P synthase are claimed, and
CC are used for the expression of the enzyme in host microbial strains,
CC especially E. coli, for use in the production of glucosamine. The
CC invention also provides methods for the overproduction of glucosamine
CC using a genetically engineered microorganism that encodes a GlcN6P
CC synthase modified to increase its activity (see also AAY58823-27).
CC Preferred modifications comprise amino acid deletion, insertion,

CC inversion, derivatisation or substitution, especially 14T, I272T, S240P,
CC A39T, R250C, G472S and L469P substitution, of the present sequence
XX
SQ Sequence 609 AA;

Query Match 100.0%; Score 3086; DB 3; Length 609;
Best Local Similarity 100.0%; Pred. No. 4.7e-264;
Matches 609; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MCGIVGAIAQRDVAEILLEGRLREYRGYDSAGLAVVDAEGHMTLRRLGKVQMLAQAAE 60
Db 1 MCGIVGAIAQRDVAEILLEGRLREYRGYDSAGLAVVDAEGHMTLRRLGKVQMLAQAAE 60
Qy 61 EPHLHGTTGIAHTRWATHGEPSEVNAHPHVSEHVVVHNGIIEHHEPLREELKARGYTFV 120
Db 61 EPHLHGTTGIAHTRWATHGEPSEVNAHPHVSEHVVVHNGIIEHHEPLREELKARGYTFV 120
Qy 121 SETDTEVIAHLVNNELKQGGLTREAVALRAIPQLRGAYGTVMDSRHPDTLLAARSGSPLV 180
Db 121 SETDTEVIAHLVNNELKQGGLTREAVALRAIPQLRGAYGTVMDSRHPDTLLAARSGSPLV 180
Qy 181 IGLGMGENFIASDQALLPVTTRFIFLEEGDIAEITRRSVNIFDKTGAEVKRODIESNLQ 240
Db 181 IGLGMGENFIASDQALLPVTTRFIFLEEGDIAEITRRSVNIFDKTGAEVKRODIESNLQ 240
Qy 241 YDAGDKGIYRHYMQEYEQPNAIKNTLTGRISHGQVDLSLGNADLLSKVEHIQILA 300
Db 241 YDAGDKGIYRHYMQEYEQPNAIKNTLTGRISHGQVDLSLGNADLLSKVEHIQILA 300
Qy 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASFRYKSAVRNLSMITLSQSGETADTLAGL 360
Db 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASFRYKSAVRNLSMITLSQSGETADTLAGL 360
Qy 361 RLSKELGYLGLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFTTQTLVLLMLVAKL 420
Db 361 RLSKELGYLGLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFTTQTLVLLMLVAKL 420
Qy 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMLSQDKRIEALAEDFSCKHHLFLGRGDQYPTA 480
Db 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMLSQDKRIEALAEDFSCKHHLFLGRGDQYPTA 480
Qy 481 LEGALKLKEISYTHABAYAAAGELKHGFLALIDAMPVIVVAPNNELLEKLKSNIEVRAR 540
Db 481 LEGALKLKEISYTHABAYAAAGELKHGFLALIDAMPVIVVAPNNELLEKLKSNIEVRAR 540
Qy 541 GGQLYYFADQDAGFVSSDNNHIIEMPHVEVIAPIETVPLQLLAYHVALIKGTDVQDPR 600
Db 541 GGQLYYFADQDAGFVSSDNNHIIEMPHVEVIAPIETVPLQLLAYHVALIKGTDVQDPR 600
Qy 601 NLAQSVTVE 609
Db 601 NLAQSVTVE 609

RESULT 2
AAU34806
ID AAU34806 standard; protein; 609 AA.
XX
AC AAU34806;
XX
DT 14-FEB-2002 (first entry)
XX
DE E. coli cellular proliferation protein #387.
XX
KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
KW antibacterial; drug design.
XX
OS Escherichia coli.
XX
PN W0200170955-A2.
XX
PD 27-SEP-2001.
XX

PF 21-MAR-2001; 2001WO-US009180..
XX
PR 21-MAR-2000; 2000US-0191078P.
PR 23-MAY-2000; 2000US-0206848P.
PR 26-MAY-2000; 2000US-0207272P.
PR 23-OCT-2000; 2000US-0242578P.
PR 27-NOV-2000; 2000US-0253625P.
PR 22-DEC-2000; 2000US-0257931P.
PR 16-FEB-2001; 2001US-0269308P.
XX (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX WPI; 2001-611495/70.
DR N-PSDB; AAS52665.
DR
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
XX
PS Example 3; SEQ ID NO 10399; Silpp; English.
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes
CC themselves and the encoded proteins. The prokaryotes used are Escherichia
CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence represents an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 609 AA;

Query Match 100.0%; Score 3086; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 4.7e-264;
Matches 609; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MCGIVGAIAQRDVAEILLEGRLREYRGYDSAGLAVVDAEGHMTLRRLGKVQMLAQAAE 60
Db 1 MCGIVGAIAQRDVAEILLEGRLREYRGYDSAGLAVVDAEGHMTLRRLGKVQMLAQAAE 60
Qy 61 EPHLHGTTGIAHTRWATHGEPSEVNAHPHVSEHVVVHNGIIEHHEPLREELKARGYTFV 120
Db 61 EPHLHGTTGIAHTRWATHGEPSEVNAHPHVSEHVVVHNGIIEHHEPLREELKARGYTFV 120
Qy 121 SETDTEVIAHLVNNELKQGGLTREAVALRAIPQLRGAYGTVMDSRHPDTLLAARSGSPLV 180
Db 121 SETDTEVIAHLVNNELKQGGLTREAVALRAIPQLRGAYGTVMDSRHPDTLLAARSGSPLV 180
Qy 181 IGLGMGENFIASDQALLPVTTRFIFLEEGDIAEITRRSVNIFDKTGAEVKRODIESNLQ 240
Db 181 IGLGMGENFIASDQALLPVTTRFIFLEEGDIAEITRRSVNIFDKTGAEVKRODIESNLQ 240
Qy 241 YDAGDKGIYRHYMQEYEQPNAIKNTLTGRISHGQVDLSLGNADLLSKVEHIQILA 300
Db 241 YDAGDKGIYRHYMQEYEQPNAIKNTLTGRISHGQVDLSLGNADLLSKVEHIQILA 300
Qy 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASFRYKSAVRNLSMITLSQSGETADTLAGL 360
Db 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASFRYKSAVRNLSMITLSQSGETADTLAGL 360
Qy 361 RLSKELGYLGLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFTTQTLVLLMLVAKL 420

Db 361 RLSKELGYLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFTTQLTVLLMLVAKL 420
Qy 421 SRLKGLDASIEHDIVHGLQALPSRIEQMLSDQKRIEALAEFSDKHHALFLGRGDQYPIA 480
Db 421 SRLKGLDASIEHDIVHGLQALPSRIEQMLSDQKRIEALAEFSDKHHALFLGRGDQYPIA 480
Qy 481 LEGALKLKEISYTHAEAYAAAGELKHGPLALIDADMPVIVVAPNNLEKLSNIEEVRAR 540
Db 481 LEGALKLKEISYTHAEAYAAAGELKHGPLALIDADMPVIVVAPNNLEKLSNIEEVRAR 540
Qy 541 GGQLYVPADQDAGFVSSDNNHIIEMPHVEVIAPIFTVTPQLLAYHVALIKGTVDVQPR 600
Db 541 GGQLYVPADQDAGFVSSDNNHIIEMPHVEVIAPIFTVTPQLLAYHVALIKGTVDVQPR 600
Qy 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 3
ABU28829
ID ABU28829 standard; protein; 609 AA.
AC ABU28829;
XX
XX 19-JUN-2003 (first entry)
XX Protein encoded by Prokaryotic essential gene #14356.
DE Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX Escherichia coli.
XX WO200277183-A2.
XX 03-OCT-2002.
XX 21-MAR-2002; 2002WO-US009107.
XX 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX (ELIT-) ELITRA PHARM INC.
PA Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
PI WPI; 2003-029926/02.
DR N-PSDB; ACA32699.
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX Claim 25; SEQ ID NO 56753; 1766pp; English.
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)

CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 609 AA;
Query Match 100.0%; Score 3086; DB 6; Length 609;
Best Local Similarity 100.0%; Pred. No. 4.7e-264;
Matches 609; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MCGIVGAIQORDVAEILLEGRLRLEYRGYDSAGLAVVDAEGHMTLRRLRGKVQMLAQAAE 60
Db 1 MCGIVGAIQORDVAEILLEGRLRLEYRGYDSAGLAVVDAEGHMTLRRLRGKVQMLAQAAE 60
Qy 61 EHPHGGTGIAHTRWATHGEPSEVNAHPHVSEHIVVHNGIIEHNPRLREELKARGYTFV 120
Db 61 EHPHGGTGIAHTRWATHGEPSEVNAHPHVSEHIVVHNGIIEHNPRLREELKARGYTFV 120
Qy 121 SETDTEVIAHLNVNWLKQGGTLREAVLRAIPQIRGAYGTVMDSRHPDITLLAARSGSPLV 180
Db 121 SETDTEVIAHLNVNWLKQGGTLREAVLRAIPQIRGAYGTVMDSRHPDITLLAARSGSPLV 180
Qy 181 IGLGMENFIASDQALLPVTRRFIFLEEGDIAEITRRSVNI FDKTGAEVKRODIESNLQ 240
Db 181 IGLGMENFIASDQALLPVTRRFIFLEEGDIAEITRRSVNI FDKTGAEVKRODIESNLQ 240
Qy 241 YDAGDKGIYRHYMQKEIYEOPNAIKNTLTGRISHGQVDLSELGNADLLELSKVEHIQILA 300
Db 241 YDAGDKGIYRHYMQKEIYEOPNAIKNTLTGRISHGQVDLSELGNADLLELSKVEHIQILA 300
Qy 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASFEFRYKSAVRNLSMITLSQSGETADTLA 360
Db 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASFEFRYKSAVRNLSMITLSQSGETADTLA 360
Qy 361 RLSKELGYLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFTTQLTVLLMLVAKL 420
Db 361 RLSKELGYLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFTTQLTVLLMLVAKL 420
Qy 421 SRLKGLDASIEHDIVHGLQALPSRIEQMLSDQKRIEALAEFSDKHHALFLGRGDQYPIA 480
Db 421 SRLKGLDASIEHDIVHGLQALPSRIEQMLSDQKRIEALAEFSDKHHALFLGRGDQYPIA 480
Qy 481 LEGALKLKEISYTHAEAYAAAGELKHGPLALIDADMPVIVVAPNNLEKLSNIEEVRAR 540
Db 481 LEGALKLKEISYTHAEAYAAAGELKHGPLALIDADMPVIVVAPNNLEKLSNIEEVRAR 540
Qy 541 GGQLYVPADQDAGFVSSDNNHIIEMPHVEVIAPIFTVTPQLLAYHVALIKGTVDVQPR 600
Db 541 GGQLYVPADQDAGFVSSDNNHIIEMPHVEVIAPIFTVTPQLLAYHVALIKGTVDVQPR 600
Qy 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609
RESULT 4
ADI38857
ID ADI38857 standard; protein; 609 AA.

XX
AC ADI38857;
DT 15-APR-2004 (first entry)
DE Glucosamine-6-phosphate synthase, glms, SEQ ID 2.
XX
KW Glucosamine; N-acetylglucosamine; fermentation;
XX Glucosamine-6-phosphate acetyltransferase;
KW Glucosamine-6-phosphate synthase; Glucosamine-6-phosphate deaminase;
KW Glucosamine-1-phosphate N-acetyltransferase; Glucosamine-6-phosphate;
KW Glucosamine-1-phosphate; N-acetylglucosamine-1-phosphate;
XX N-acetylglucosamine-6-phosphate; enzyme.
OS Escherichia coli.
XX
XX WO2004003175-A2.
XX
XX 08-JAN-2004.
XX
XX 01-JUL-2003; 2003WO-US020925.
XX
XX 01-JUL-2002; 2002US-0393348P.
XX
XX (ARKI-) ARKION LIFE SCI LLC.
XX
XX Deng M, Angerer JD, Cyron D, Grund AD, Jerrell TA, Leanna C;
PI Mathre O, Rosson R, Running J, Severson D, Song L, Wassink S;
XX
XX WPI; 2004-203380/19.
DR N-PSDB; ADI38856.
XX
XX
XX Producing microorganism or N-acetylglucosamine by fermentation involves
PT culturing microorganism comprising glucosamine-6-phosphate
PT acetyltransferase, in fermentation medium, and collecting product.
XX
XX Claim 15; SEQ ID NO 2; 327pp; English.
XX
XX The present invention relates to a method (M1) for producing glucosamine
CC and N-acetylglucosamine by fermentation. The method comprises (a)
CC culturing in a fermentation medium a microorganism (I) which comprises
CC endogenous glucosamine-6-phosphate acetyltransferase (II) and a genetic
CC modification that increases the activity of (II), glucosamine-6-phosphate
CC synthase (III) or glucosamine-6-phosphate deaminase (IV), or decreases
CC the activity of (IV) and increases the activity of glucosamine-1-
CC phosphate N-acetyltransferase (V), and (b) and collecting the product,
CC which is chosen from the group consisting of glucosamine-6-phosphate,
CC glucosamine, glucosamine-1-phosphate, N-acetylglucosamine-1-phosphate,
CC acetylglucosamine-6-phosphate, and N-acetylglucosamine. The present
CC sequence was used to illustrate the method of the invention.
XX
SQ Sequence 609 AA;

Query Match 100.0%; Score 3086; DB 8; Length 609;
Best Local Similarity 100.0%; Pred. No. 4.7e-264;
Matches 609; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGIVGAIARDVAEILLEGRLRLEVRGYDSAGLAVVDAEGHMTLRRLRGKVOMLAQAAE 60
DB 1 MCGIVGAIARDVAEILLEGRLRLEVRGYDSAGLAVVDAEGHMTLRRLRGKVOMLAQAAE 60
QY 61 EHPHGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIENHEPLREELKARGYTFV 120
DB 61 EHPHGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIENHEPLREELKARGYTFV 120
QY 121 SETDTEVIAHLNVNWLKQGTLEAVLRAIPQIRGAYGTIVMDSRHPDTLLAARSGSLV 180
DB 121 SETDTEVIAHLNVNWLKQGTLEAVLRAIPQIRGAYGTIVMDSRHPDTLLAARSGSLV 180
QY 181 IGLGMENFIASDQALLPVTRFRFLEEGDIAETTRRSVNI FDKTGAEVKQDIESNLQ 240
DB 181 IGLGMENFIASDQALLPVTRFRFLEEGDIAETTRRSVNI FDKTGAEVKQDIESNLQ 240

QY 241 YDAGDKGIYRHYMQKEIYEQPNAIKNTLTGRISHGQVDLSELGPNADLLSKVEHIQILA 300
DB 241 YDAGDKGIYRHYMQKEIYEQPNAIKNTLTGRISHGQVDLSELGPNADLLSKVEHIQILA 300
QY 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASFEPRYKSAVRNLSMITLSQSGTADTLAGL 360
DB 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASFEPRYKSAVRNLSMITLSQSGTADTLAGL 360
QY 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 420
DB 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 420
QY 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMLSQDKRIEALAEFSDKHAFFLGRGQDYPFA 480
DB 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMLSQDKRIEALAEFSDKHAFFLGRGQDYPFA 480
QY 481 LEGALKLKEISYIHAEEAAGELKHGPLALIDAMPVIVVAPNNELLEKLKSNIEEVRAR 540
DB 481 LEGALKLKEISYIHAEEAAGELKHGPLALIDAMPVIVVAPNNELLEKLKSNIEEVRAR 540
QY 541 GGQLYVFADQDAGFVSSDNMHIIEMPHVEEVIPIFYTVPLQLLAYHVALIKGTDVDQPR 600
DB 541 GGQLYVFADQDAGFVSSDNMHIIEMPHVEEVIPIFYTVPLQLLAYHVALIKGTDVDQPR 600
QY 601 NLAKSVTVVE 609
DB 601 NLAKSVTVVE 609

RESULT 5
ADS45181
ID ADS45181 standard; protein; 609 AA.
XX
AC ADS45181;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #23611.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
OS Bacteria.
XX
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
XX (CAOY/) CAO Y.
XX (HINK/) HINKLE G J.
XX (SLAT/) SLATER S C.
XX (CHEN/) CHEN X.
XX (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 23611; 122pp; English.
XX

CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX SQ Sequence 609 AA;

Query Match 100.0%; Score 3086; DB 8; Length 609;
Best Local Similarity 100.0%; Pred. No. 4.7e-264;
Matches 609; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGIVGAIARDVAEILLEGRLREYRGYSAGLAVVDAEGHMTLRRLRGKVMQAAAE 60
DB 1 MCGIVGAIARDVAEILLEGRLREYRGYSAGLAVVDAEGHMTLRRLRGKVMQAAAE 60
QY 61 EHPLHGCGTGAHTRWATHGPSEVNAHPVSEHIVVHNGIENHEPLREELKARGYTFV 120
DB 61 EHPLHGCGTGAHTRWATHGPSEVNAHPVSEHIVVHNGIENHEPLREELKARGYTFV 120
QY 121 SETDTEVIAHVNWELKOGGTLRAVLRAIPOLRGAYGTWIMDSRHPDTHLAARSGSPLV 180
DB 121 SETDTEVIAHVNWELKOGGTLRAVLRAIPOLRGAYGTWIMDSRHPDTHLAARSGSPLV 180
QY 181 IGLGMGNFTASDQALLPVTRRIFLEEGDIAETRISVNIPOKTAEGVKRQDIESNLQ 240
DB 181 IGLGMGNFTASDQALLPVTRRIFLEEGDIAETRISVNIPOKTAEGVKRQDIESNLQ 240
QY 241 YDAGDKGIYRHYMQEYEQPNAIKNTLTGRISHGQVDLSLGNADLLSKVEHIQILA 300
DB 241 YDAGDKGIYRHYMQEYEQPNAIKNTLTGRISHGQVDLSLGNADLLSKVEHIQILA 300
QY 301 CGTSYNGMYSRYWFESLAGIPCDVEIASBFRYKSAVRNSLMITLSQSGETADTLA 360
DB 301 CGTSYNGMYSRYWFESLAGIPCDVEIASBFRYKSAVRNSLMITLSQSGETADTLA 360
QY 361 RLSKELGYLSLAI CNVPGSSLVRESLALMTNAGTIGVASTKAFQTTLVLLMLVAKL 420
DB 361 RLSKELGYLSLAI CNVPGSSLVRESLALMTNAGTIGVASTKAFQTTLVLLMLVAKL 420
QY 421 SRLKGLDASIEHDIVHGLQALPSRIEQLMSQDKRIEALAEFDSKHHALFLGRGDQYPIA 480
DB 421 SRLKGLDASIEHDIVHGLQALPSRIEQLMSQDKRIEALAEFDSKHHALFLGRGDQYPIA 480
QY 481 LEGALKLKEISYTHAEAYAGELKHGFLALIDAMPVIVVAPNNLEKLSNIEEVRAR 540
DB 481 LEGALKLKEISYTHAEAYAGELKHGFLALIDAMPVIVVAPNNLEKLSNIEEVRAR 540
QY 541 GGQLYVFADQDAGFVSSDNHIIEMPHVEEVIAPFVTPQLLIAYHVALIKGTDVDDQPR 600
DB 541 GGQLYVFADQDAGFVSSDNHIIEMPHVEEVIAPFVTPQLLIAYHVALIKGTDVDDQPR 600
QY 601 NLAKSVTVE 609
|||||

Db 601 NLAKSVTVE 609

RESULT 6

ID ADU00590
XX ADU00590 standard; protein; 609 AA.

AC ADU00590;

DT 13-JAN-2005 (first entry)

XX Amino acid sequence of a glucosamine-6-phosphate synthase.

XX chitin; chitosan; fermentation;
KW glutamine-fructose-6-phosphate amidotransferase; glucosamine synthetase;
KW glucosamine-6-phosphate synthase; GPAL gene; glms gene;
KW glucosamine-6-phosphate acetyltransferase; GNA1 gene; chitin synthase;
KW chitin deacetylase; CDAL gene; CDA2 gene;
KW N-acetylglucosamine-6-phosphate deacetylase;
KW glucosamine-6-phosphate deaminase; chitinase; chitosanase; fungus; yeast;
KW enzyme.

OS Escherichia coli.

XX WO2004092391-A2.

XX 28-OCT-2004.

XX 12-APR-2004; 2004WO-US011286.

XX 11-APR-2003; 2003US-0462087P.

XX (ARKI-) ARKION LIFE SCI LLC.

PI Deng M, McMullin TW, Grund AD;

XX WPI; 2004-766880/75.

XX N-PSDB; ADU00589.

XX Producing chitin or chitosan, involves culturing microorganism with
PT genetic modification that results in increase in activity of glutamine-
PT fructos-6-phosphate amidotransferase, in fermentation medium, to produce
PT chitin or chitosan.

XX Example 1; SEQ ID NO 23; 161pp; English.

XX The specification describes a method for producing chitin or chitosan by
CC a fermentation process. The method involves culturing in a fermentation
CC medium a microorganism which comprises one or more genetic modifications
CC that result in an increase in the activity of glutamine-fructose-6-
CC phosphate amidotransferase (also known as glucosamine synthetase and
CC glucosamine-6-phosphate synthase, and encoded by the GPAL eukaryotic gene
CC and the glms bacterial gene), glucosamine-6-phosphate acetyltransferase
CC (encoded by GNA1), chitin synthase or chitin deacetylase (encoded by CDAL
CC and CDA2), or in decrease in the activity of N-acetylglucosamine-6-
CC phosphate deacetylase, glucosamine-6-phosphate deaminase, chitinase and
CC chiteanase, and collecting the chitin or chitosan. The method is useful
CC for producing chitin or chitosan by utilizing microorganisms such as
CC fungus, yeast (e.g. Saccharomyces or Schizosaccharomyces) and
CC filamentous fungus (e.g. Aspergillus, Absidia or Rhizopus), preferably S.
CC cerevisiae, A. niger or A. nidulans. The method enables high quantities
CC of chitin and chitosan to be produced cost effectively. The present
CC sequence represents an Escherichia coli glucosamine-6-phosphate synthase.
CC It was used to transform yeast for use in the method of the invention.

XX SQ Sequence 609 AA;

Query Match 100.0%; Score 3086; DB 8; Length 609;

Best Local Similarity 100.0%; Pred. No. 4.7e-264;

Matches 609; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGIVGAIARDVAEILLEGRLREYRGYSAGLAVVDAEGHMTLRRLRGKVMQAAAE 60
|||||

| | | | | | | | | | | | | | | | | | | |
|----|-----|----------|--------|----|-------|-------|-----|-----|----|-------|----|----|-----|----|----|----|-----|-----|
| Db | 1 | MCGIVGAI | AQRDAE | II | LEGLR | BEYRG | YDS | AGL | AV | DAEGH | MT | RL | BLG | VQ | ML | QA | AAE | 60 |
| Qy | 61 | EHP | LHG | GT | G | I | A | T | R | W | A | T | H | E | P | S | B | 120 |
| Db | 61 | EHP | LHG | GT | G | I | A | T | R | W | A | T | H | E | P | S | B | 120 |
| Qy | 121 | SET | TE | VI | A | H | L | V | N | W | E | L | K | Q | G | T | L | 180 |
| Db | 121 | SET | TE | VI | A | H | L | V | N | W | E | L | K | Q | G | T | L | 180 |
| Qy | 181 | I | G | L | G | M | G | N | F | T | A | S | D | Q | L | A | L | 240 |
| Db | 181 | I | G | L | G | M | G | N | F | T | A | S | D | Q | L | A | L | 240 |
| Qy | 241 | Y | D | A | G | K | I | Y | R | H | Y | M | K | E | I | Y | E | 300 |
| Db | 241 | Y | D | A | G | K | I | Y | R | H | Y | M | K | E | I | Y | E | 300 |
| Qy | 301 | C | G | T | S | Y | N | G | M | V | S | R | Y | F | E | S | L | 360 |
| Db | 301 | C | G | T | S | Y | N | G | M | V | S | R | Y | F | E | S | L | 360 |
| Qy | 361 | R | L | S | K | E | L | G | L | G | S | L | A | I | C | N | P | 420 |
| Db | 361 | R | L | S | K | E | L | G | L | G | S | L | A | I | C | N | P | 420 |
| Qy | 421 | S | R | L | K | G | L | D | A | S | I | E | H | D | I | V | H | 480 |
| Db | 421 | S | R | L | K | G | L | D | A | S | I | E | H | D | I | V | H | 480 |
| Qy | 481 | L | E | G | A | L | K | E | I | S | I | H | A | E | A | Y | A | 540 |
| Db | 481 | L | E | G | A | L | K | E | I | S | I | H | A | E | A | Y | A | 540 |
| Qy | 541 | G | G | Q | Y | F | A | D | Q | A | G | F | V | S | S | D | N | 600 |
| Db | 541 | G | G | Q | Y | F | A | D | Q | A | G | F | V | S | S | D | N | 600 |
| Qy | 601 | N | L | A | K | S | V | T | E | | | | | | | | | 609 |
| Db | 601 | N | L | A | K | S | V | T | E | | | | | | | | | 609 |

| | |
|----------|--|
| RESULT 7 | |
| ADW23840 | |
| ID | ADW23840 standard; protein; 608 AA. |
| XX | |
| XX | ADW23840; |
| AC | |
| XX | |
| XX | 07-APR-2005 (first entry) |
| XX | |
| DE | Novel human GFAT protein-related E coli GFAT protein SeqID13. |
| XX | |
| XX | protein purification; glutamine:fructose-6-phosphate amidotransferase; |
| KW | GFAT; antidiabetic; anorectic; metabolic; cytostatic; osteopathic; |
| KW | fungicide; herbicide; diabetes; non-insulin dependent diabetes; obesity; |
| KW | acidosis; cancer; osteoporosis. |

XX New modified glutamine:fructose-6-phosphate amidotransferase, useful for
PT identifying specific inhibitors, potentially useful for treating e.g.
PT diabetes, contains a purification tag, also related nucleic acid.
XX
PS Disclosure; SEQ ID NO 13; 61pp; French.
XX
CC This invention relates to a novel protein corresponding to human
CC enzymatically active but modified glutamine:fructose-6-phosphate
CC amidotransferase (GFAT). The invention may be useful for the production
CC of compounds with an antidiabetic, anorectic, metabolic, cytostatic,
CC osteopathic, fungicide or herbicide activity. The invention may be used
CC to screen for agents that modify, especially inhibit, its activity,
CC potentially useful as agents for treatment or prevention of diabetes
CC (particularly type II), obesity, acidosis, ketosis, cancer and
CC osteoporosis. In addition inhibitors of plant and fungal GFAT may be
CC useful as herbicides and fungicides, respectively. The present sequence
CC is that of a human GFAT-like protein of the invention.

| | | |
|-----|---------------------------|--|
| Seq | Sequence | 608 AA; |
| | Query Match | 99.8%; Score 3081; DB 9; Length 608; |
| | Best Local Similarity | 100.0%; Pred. No. 1.3e-263; |
| | Matches 608; Conservative | 0; Mismatches 0; Indels 0; Gaps |
| QY | 2 | CGIVGAIAORDVAEILLEGRLREYRGVDSAGLAVVDAGCHWTRLRLGKVKOMLAQAEE 61 |
| DB | 1 | CGIVGAIAORDVAEILLEGRLREYRGVDSAGLAVVDAGCHWTRLRLGKVKOMLAQAEE 60 |
| QY | 62 | HPLHGCTGIATRWATHGSEPVNAHPHVSBIHVVVHNGIIEENHEPULREELKARGYTFFVS 121 |
| DB | 61 | HPLHGCTGIATRWATHGSEPVNAHPHVSBIHVVVHNGIIEENHEPULREELKARGYTFFVS 120 |
| QY | 122 | ETDTEVIHLVNVELKQGGLTBEAFLRAIPOLRGAYGTVMDSRPHDTLLAAKSGSPIVI 181 |
| DB | 121 | ETDTEVIHLVNVELKQGGLTBEAFLRAIPOLRGAYGTVMDSRPHDTLLAAKSGSPIVI 180 |
| QY | 182 | GLGMGNFIASDLQALLPVTRRFIFLEEGDIAETRRSVNIFDKTGAEVRQDIIESNLQY 241 |
| DB | 181 | GLGMGNFIASDLQALLPVTRRFIFLEEGDIAETRRSVNIFDKTGAEVRQDIIESNLQY 240 |
| QY | 242 | DAGDGKIYRHYNQKEIIYEQPNAINKTLTGRISHQGVDSLSELGPNADELLSKVEHIQILAC 301 |
| DB | 241 | DAGDGKIYRHYNQKEIIYEQPNAINKTLTGRISHQGVDSLSELGPNADELLSKVEHIQILAC 300 |
| QY | 302 | GTSYNSGMVSRYPWFESLAGICPDVEIASEFPRYSKSAVRNLSMTITLSQSETADTLA GLR 361 |
| DB | 301 | GTSYNSGMVSRYPWFESLAGICPDVEIASEFPRYSKSAVRNLSMTITLSQSETADTLA GLR 360 |
| QY | 362 | LSEKLGYLGSLAICNVPOSSIVRESDLALMTNAGTEIGVASTKAFTTQLTVLLMLVAKLS 421 |
| DB | 361 | LSEKLGYLGSLAICNVPOSSIVRESDLALMTNAGTEIGVASTKAFTTQLTVLLMLVAKLS 420 |
| QY | 422 | RUKGLDASIETHDIVHGLQALPSRIEOMLSODKRJEALAEPDFSKHHALFLGRGDQPIAL 481 |
| DB | 421 | RUKGLDASIETHDIVHGLQALPSRIEOMLSODKRJEALAEPDFSKHHALFLGRGDQPIAL 480 |
| QY | 482 | EGALKLKEISITHAEAYAAGBLKHGPIALIDAMPVIVVAPNNELLEKKSNTEEVRAR 541 |
| DB | 481 | EGALKLKEISITHAEAYAAGBLKHGPIALIDAMPVIVVAPNNELLEKKSNTEEVRAR 540 |
| QY | 542 | GQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPIFYTPVLQLLAYHVALIKGTDVDQPRN 601 |
| DB | 541 | GQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPIFYTPVLQLLAYHVALIKGTDVDQPRN 600 |
| QY | 602 | LAKSVTVE 609 |
| DB | 601 | LAKSVTVE 608 |

RESULT 8
AAY58827
ID AAY58827 standard; protein; 609 AA.

CC coli, that includes a modified Glc6P synthase. Production of the
CC glucosamine by recombinant strain 2123-149 (pKLN23-149) was marginally
CC increased when compared to a strain expressing wild-type Glc6P synthase
XX
SQ Sequence 609 AA;

Query Match 99.8%; Score 3080; DB 3; Length 609;
Best Local Similarity 99.8%; Pred. No. 1.6e-263;
Matches 608; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MCGIVGAIARDVAEILLEGRLRLEYGYDSAGLAVVDAEGHMTLRRLRGKQVQLAAAE 60
Db 1 MCGIVGAIARDVAEILLEGRLRLEYGYDSAGLAVVDAEGHMTLRRLRGKQVQLAAAE 60
Qy 61 EPHLHGCGTGAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHPELREELKARGYTFV 120
Db 61 EPHLHGCGTGAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHPELREELKARGYTFV 120
Qy 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPV 180
Db 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPV 180
Qy 181 IGLGMGFASDQALLPVTRRFIFLEEGDIAEITRRSVNIPDKTGAEVKQDIESNLQ 240
Db 181 IGLGMGFASDQALLPVTRRFIFLEEGDIAEITRRSVNIPDKTGAEVKQDIESNLQ 240
Qy 241 YDAGDKGIYRHYMKEIYEOPNAIKNTLTGRISHGQVDLSLPGNADELLSKVEHIQILA 300
Db 241 YDAGDKGIYRHYMKEIYEOPNAIKNTLTGRISHGQVDLSLPGNADELLSKVEHIQILA 300
Qy 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASEFRYKSAVRNLSMITLSQGETADTLA 360
Db 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASEFRYKSAVRNLSMITLSQGETADTLA 360
Qy 361 RLSKELGYLGSIAICNVPSSLVRESDLALMTNAGTEIGVASTKAFTTQLTVLLMLVAKL 420
Db 361 RLSKELGYLGSIAICNVPSSLVRESDLALMTNAGTEIGVASTKAFTTQLTVLLMLVAKL 420
Qy 421 SRLKGLDASIEHDIVHGLQALPSRIEQMLSODKRIEALAEFSDKHAFILGRGDQYPTA 480
Db 421 SRLKGLDASIEHDIVHGLQALPSRIEQMLSODKRIEALAEFSDKHAFILGRGDQYPTA 480
Qy 481 LEGALKKEISYTHAAYAAGELKHGFLALIDAMPVIVVAPNNELLEKLKNIIEVRAR 540
Db 481 LEGALKKEISYTHAAYAAGELKHGFLALIDAMPVIVVAPNNELLEKLKNIIEVRAR 540
Qy 541 GGQLYVFADQDAGFVSSDNNHIIEMPHVEVIAPIFYTVPLQLLAYVHALIKGTDVDP 600
Db 541 GGQLYVFADQDAGFVSSDNNHIIEMPHVEVIAPIFYTVPLQLLAYVHALIKGTDVDP 600
Qy 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 10
AD138865
ID AD138865 standard; protein; 609 AA.
XX
AC AD138865;
XX
DT 15-APR-2004 (first entry)
XX
DE Mutant glucosamine-6-phosphate synthase, SEQ ID 10.
XX
KW Glucosamine; N-acetylglucosamine; fermentation;
KW glucosamine-6-phosphate acetyltransferase;
KW glucosamine-6-phosphate synthase; glucosamine-6-phosphate deaminase;
KW glucosamine-1 phosphate N-acetyltransferase; glucosamine-6-phosphate;
KW glucosamine-1-phosphate; N-acetylglucosamine-1-phosphate;
XX N-acetylglucosamine-6-phosphate; enzyme.
XX Escherichia coli.

OS Synthetic.
XX WO2004003175-A2.
PN
XX 08-JAN-2004.
PD
PF 01-JUL-2003; 2003WO-US020925.
XX
PR 01-JUL-2002; 2002US-0393348P.
XX
PA (ARKI-) ARKION LIFE SCI LLC.
XX
PI Deng M, Angerer JD, Cyron D, Grund AD, Jerrell TA, Leanna C;
PI Mathre O, Rosson R, Running J, Severson D, Song L, Wassink S;
XX
DR WPI; 2004-203380/19.
DR N-PSDB; ADI38864.
XX
PT Producing glucosamine or N-acetylglucosamine by fermentation involves
PT culturing microorganism comprising glucosamine-6-phosphate
PT acetyltransferase, in fermentation medium, and collecting product.
XX
PS Claim 15; SEQ ID NO 10; 327pp; English.
XX
CC The present invention relates to a method (M1) for producing glucosamine
CC and N-acetylglucosamine by fermentation. The method comprises (a)
CC culturing in a fermentation medium a microorganism (I) which comprises
CC endogenous glucosamine-6-phosphate acetyltransferase (II) and a genetic
CC modification that increases the activity of (II), glucosamine-6-phosphate
CC synthase (III) or glucosamine-6-phosphate deaminase (IV), or decreases
CC the activity of (IV) and increases the activity of glucosamine-1-
CC phosphate N-acetyltransferase (V), and (b) and collecting the product,
CC which is chosen from the group consisting of glucosamine-6-phosphate, N-
CC glucosamine, glucosamine-1-phosphate, N-acetylglucosamine-1-phosphate, N-
CC acetylglucosamine-6-phosphate, and N-acetylglucosamine. The present
CC sequence was used to illustrate the method of the invention.
XX
SQ Sequence 609 AA;

Query Match 99.8%; Score 3080; DB 8; Length 609;
Best Local Similarity 99.8%; Pred. No. 1.6e-263;
Matches 608; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MCGIVGAIARDVAEILLEGRLRLEYGYDSAGLAVVDAEGHMTLRRLRGKQVQLAAAE 60
Db 1 MCGIVGAIARDVAEILLEGRLRLEYGYDSAGLAVVDAEGHMTLRRLRGKQVQLAAAE 60
Qy 61 EPHLHGCGTGAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHPELREELKARGYTFV 120
Db 61 EPHLHGCGTGAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHPELREELKARGYTFV 120
Qy 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPV 180
Db 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPV 180
Qy 181 IGLGMGFASDQALLPVTRRFIFLEEGDIAEITRRSVNIPDKTGAEVKQDIESNLQ 240
Db 181 IGLGMGFASDQALLPVTRRFIFLEEGDIAEITRRSVNIPDKTGAEVKQDIESNLQ 240
Qy 241 YDAGDKGIYRHYMKEIYEOPNAIKNTLTGRISHGQVDLSLPGNADELLSKVEHIQILA 300
Db 241 YDAGDKGIYRHYMKEIYEOPNAIKNTLTGRISHGQVDLSLPGNADELLSKVEHIQILA 300
Qy 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASEFRYKSAVRNLSMITLSQGETADTLA 360
Db 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASEFRYKSAVRNLSMITLSQGETADTLA 360
Qy 361 RLSKELGYLGSIAICNVPSSLVRESDLALMTNAGTEIGVASTKAFTTQLTVLLMLVAKL 420
Db 361 RLSKELGYLGSIAICNVPSSLVRESDLALMTNAGTEIGVASTKAFTTQLTVLLMLVAKL 420
Qy 421 SRLKGLDASIEHDIVHGLQALPSRIEQMLSODKRIEALAEFSDKHAFILGRGDQYPTA 480

Db 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMLSDQKRIEALAEFSDKHAFLSRGDQVPIA 480
Qy 481 LEGALKLKEISYTHAEAYAGELKHGPLALIDAMPVIVVAPNNLEKLSNTEEVRR 540
Db 481 LEGALKLKEISYTHAEAYAGELKHGPLALIDAMPVIVVAPNNLEKLSNTEEVRR 540
Qy 541 GGQLYVPADQDAGFVSSDNMHIIMPHVEEVIPIFYTVPLQLLAYHVALIKGTDVDPQR 600
Db 541 GGQLYVPADQDAGFVSSDNMHIIMPHVEEVIPIFYTVPLQLLAYHVALIKGTDVDPQR 600
Qy 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 11
ADI38867
ID ADI38867 standard; protein; 609 AA.

XX ADI38867;
XX 15-APR-2004 (first entry)
XX Mutant glucosamine-6-phosphate synthase, SEQ ID 12.
XX Glucosamine; N-acetylglucosamine; fermentation;
KW glucosamine-6-phosphate acetyltransferase;
KW glucosamine-6-phosphate synthase; glucosamine-6-phosphate deaminase;
KW glucosamine-1-phosphate N-acetyltransferase; glucosamine-6-phosphate;
KW glucosamine-1-phosphate; N-acetylglucosamine-1-phosphate;
KW N-acetylglucosamine-6-phosphate; enzyme.

XX Escherichia coli.
OS Synthetic.
XX WO2004003175-A2.
XX 08-JAN-2004.
XX 01-JUL-2003; 2003WO-US020925.
XX 01-JUL-2002; 2002US-0393348P.
XX (ARKI-) ARKION LIFE SCI LLC.
XX Deng M, Angerer JD, Cyron D, Grund AD, Jerrell TA, Leanna C;
PI Mathre O, Rosson R, Running J, Severson D, Song L, Wassink S;
XX WPI; 2004-203380/19.
DR N-PSDB; ADI38866.

XX Producing glucosamine or N-acetylglucosamine by fermentation involves
PT culturing microorganism comprising glucosamine-6-phosphate
PT acetyltransferase, in fermentation medium, and collecting product.

XX Claim 15; SEQ ID NO 12; 327pp; English.

XX The present invention relates to a method (M1) for producing glucosamine
CC and N-acetylglucosamine by fermentation. The method comprises (a)
CC culturing in a fermentation medium a microorganism (I) which comprises
CC endogenous glucosamine-6-phosphate acetyltransferase (II) and a genetic
CC modification that increases the activity of (II), glucosamine-6-phosphate
CC synthase (III) or glucosamine-6-phosphate deaminase (IV), or decreases
CC the activity of (IV) and increases the activity of glucosamine-1
CC phosphate N-acetyltransferase (V), and (b) and collecting the product,
CC which is chosen from the group consisting of glucosamine-6-phosphate,
CC glucosamine, glucosamine-1-phosphate, N-acetylglucosamine-1-phosphate, N-
CC acetylglucosamine-6-phosphate, and N-acetylglucosamine. The present
CC sequence was used to illustrate the method of the invention.

XX Sequence 609 AA;

Query Match 99.8%; Score 3080; DB 8; Length 609;

Best Local Similarity 99.8%; Pred. No. 1.6e-263;
Matches 608; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MCGTVGAIAORDVAEIILEGRLRLEYRGYDSAGLAVVDAEGHMTLRRLRGKQVQLAAAE 60
Db 1 MCGLVGAIAORDVAEIILEGRLRLEYRGYDSAGLAVVDAEGHMTLRRLRGKQVQLAAAE 60
Qy 61 EPHLHGCTGIAHTRWATHGEPSEVNAHPHVSEHIVVHNGIIENHEPLREELKARGYTFV 120
Db 61 EPHLHGCTGIAHTRWATHGEPSEVNAHPHVSEHIVVHNGIIENHEPLREELKARGYTFV 120
Qy 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDPTLLAARSGSPLV 180
Db 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDPTLLAARSGSPLV 180
Qy 181 IGLGMENFIASDQALLPVTRRPFIEEGDIAEITRRSVNIPDKTCAEVKRDIESNIQ 240
Db 181 IGLGMENFIASDQALLPVTRRPFIEEGDIAEITRRSVNIPDKTCAEVKRDIESNIQ 240
Qy 241 YDAGDKGIYRHYMQKEIYEOPNAIKNTLTGRISHGQVDLSSELGNADLLSKVEHIQILA 300
Db 241 YDAGDKGIYRHYMQKEIYEOPNAIKNTLTGRISHGQVDLSSELGNADLLSKVEHIQILA 300
Qy 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASBFRYKSAVRNNSLMITLSQSGETADTLA 360
Db 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASBFRYKSAVRNNSLMITLSQSGETADTLA 360
Qy 361 RLSKELGYLGLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFPTTQTLVLLMLVAKL 420
Db 361 RLSKELGYLGLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFPTTQTLVLLMLVAKL 420
Qy 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMLSDQKRIEALAEFSDKHAFLSRGDQVPIA 480
Db 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMLSDQKRIEALAEFSDKHAFLSRGDQVPIA 480
Qy 481 LEGALKLKEISYTHAEAYAGELKHGPLALIDAMPVIVVAPNNLEKLSNTEEVRR 540
Db 481 LEGALKLKEISYTHAEAYAGELKHGPLALIDAMPVIVVAPNNLEKLSNTEEVRR 540
Qy 541 GGQLYVPADQDAGFVSSDNMHIIMPHVEEVIPIFYTVPLQLLAYHVALIKGTDVDPQR 600
Db 541 GGQLYVPADQDAGFVSSDNMHIIMPHVEEVIPIFYTVPLQLLAYHVALIKGTDVDPQR 600
Qy 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 12
AAVS8825

ID AAVS8825 standard; protein; 609 AA.

XX AAVS8825;

XX 08-MAY-2000 (first entry)

XX E. coli glucosamine-6-phosphate synthase mutant GlcN6P-S-124.

XX Glucosamine-6-phosphate synthase; glms gene; mutant; GlcN6P-S-124;
KW glucosamine; metabolic engineering; plasmid pKLN23-124; mutcin.

XX Escherichia coli.

XX Key Location/Qualifiers

FT Misc-difference 469 /note= "replaces wild-type Leu"

XX WO200004182-A1.

XX 27-JAN-2000.

XX 15-JUL-1999; 99WO-US015976.

PR 15-JUL-1998; 98US-00115475.
XX (DCVB-) DCV INC DBA BIO-TECH RESOURCES.
XX Berry A, Burlingame RP, Millis JR;
PI WPI; 2000-182441/16.
DR N-PSDB; AAZ58252.
XX
PT Fermentation of E. coli having an altered amino acid sugar metabolic
PT pathway to produce glucosamine, especially using novel recombinant
PT variant glucosamine-6-phosphate synthases.
XX
PS Claim 28; Page 130-132; 150pp; English.
XX
CC The present sequence is that of a mutant, denoted GlcN6P-S-124, of the
CC glucosamine-6-phosphate synthase (GlcN6P synthase) of Escherichia coli.
CC When compared with the wild-type sequence (see AAY58822), the mutant
CC includes a Leu-469 to Thr amino acid substitution. This alteration is
CC predicted from the mutated glms gene in plasmid pKLN23-124 (see AAZ58252).
CC The invention provides methods for the overproduction of glucosamine by
CC fermentation using a genetically engineered microorganism, especially E.
CC coli, that includes a modified GlcN6P synthase. Production of the
CC glucosamine by recombinant strain 2123-124 (pKLN23-124) was significantly
CC increased when compared to a strain expressing wild-type Glc6NP synthase
CC owing to reduced product inhibition
XX
SQ Sequence 609 AA;

Query Match 99.8%; Score 3079; DB 3; Length 609;
Best Local Similarity 99.8%; Pred. No. 2e-263;
Matches 608; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MCGIVGAIAQDVAEILLEGRLRYGYDSAGLAVVDAEGHMTLRRLRGKVMQAQAAE 60
Db 1 MCGIVGAIAQDVAEILLEGRLRYGYDSAGLAVVDAEGHMTLRRLRGKVMQAQAAE 60

Qy 61 EPHLHGCGTGAHTRWATHGEPSEVNAHPHVSHEHVVVHNGIIEHHEPLREELKARGYTFV 120
Db 61 EPHLHGCGTGAHTRWATHGEPSEVNAHPHVSHEHVVVHNGIIEHHEPLREELKARGYTFV 120

Qy 121 SETDTEVIAHLVNWELKQGTLEAVLRATPQLRGAYGTVMDSRHPDTLLAARSGSPVLV 180
Db 121 SETDTEVIAHLVNWELKQGTLEAVLRATPQLRGAYGTVMDSRHPDTLLAARSGSPVLV 180

Qy 181 IGLGMGNFTASDQALLPVTRRFIFLEEGDIAEITRVSNIPTDKTGAEVKQDIESNLQ 240
Db 181 IGLGMGNFTASDQALLPVTRRFIFLEEGDIAEITRVSNIPTDKTGAEVKQDIESNLQ 240

Qy 241 YDAGDKGIYRHYMQEYIYQPNNAIKNTLTGRISHGQVDLSGLPNADELLSKVEHIQILA 300
Db 241 YDAGDKGIYRHYMQEYIYQPNNAIKNTLTGRISHGQVDLSGLPNADELLSKVEHIQILA 300

Qy 301 CGTSYNSGMVSRVWFSLSAGIPCDVEIASEFRYKSAVRNSLMITLSQSGTADTLAAGL 360
Db 301 CGTSYNSGMVSRVWFSLSAGIPCDVEIASEFRYKSAVRNSLMITLSQSGTADTLAAGL 360

Qy 361 RLSKELGYLGLAICNVPSSLVRESDLALMTNAGTEIGVASTKFTTTLVLMVAKL 420
Db 361 RLSKELGYLGLAICNVPSSLVRESDLALMTNAGTEIGVASTKFTTTLVLMVAKL 420

Qy 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMLSQDKRIEALAEFSDKHAFLLGRGQYPIA 480
Db 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMLSQDKRIEALAEFSDKHAFLLGRGQYPIA 480

Qy 481 LEGALKKEISYTHAAYAGELKHGFLALIDADMPVIVVAPNNLEKLKSNIEVRAR 540
Db 481 LEGALKKEISYTHAAYAGELKHGFLALIDADMPVIVVAPNNLEKLKSNIEVRAR 540

Qy 541 GGQLYVFAQDQAGFVSSDNNHIIEMPHVEVIAPIFYTVPLQLLAYHVALIKGTDVDDQR 600
Db 541 GGQLYVFAQDQAGFVSSDNNHIIEMPHVEVIAPIFYTVPLQLLAYHVALIKGTDVDDQR 600

Qy 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 13
ADI38863
ID ADI38863 standard; protein; 609 AA.
XX
XX ADI38863;
AC
XX 15-APR-2004 (first entry)
DT
XX
DE Mutant glucosamine-6-phosphate synthase, glms*124, SEQ ID 8.
XX
XX Glucosamine; N-acetylglucosamine; fermentation;
KW glucosamine-6-phosphate acetyltransferase;
KW glucosamine-6-phosphate synthase; glucosamine-6-phosphate deaminase;
KW glucosamine-1 phosphate N-acetyltransferase; glucosamine-6-phosphate;
KW glucosamine-1-phosphate; N-acetylglucosamine-1-phosphate;
KW N-acetylglucosamine-6-phosphate; enzyme.
XX
OS Escherichia coli.
OS Synthetic.
XX
XX WO2004003175-A2.
XX
XX 08-JAN-2004.
PD
XX
XX 01-JUL-2003; 2003WO-US020925.
XX
XX 01-JUL-2002; 2002US-0393348P.
PR
XX (ARKI-) ARKION LIFE SCI LLC.
PA
XX Deng M, Angerer JD, Cyron D, Grund AD, Jerrell TA, Leanna C;
PI Mathre O, Rosson R, Running J, Severson D, Song L, Wassink S;
XX
XX WPI; 2004-203380/19.
DR
DR N-PSDB; ADI38862.
XX
XX
PT Producing glucosamine or N-acetylglucosamine by fermentation involves
PT culturing microorganism comprising glucosamine-6-phosphate
PT acetyltransferase, in fermentation medium, and collecting product.
XX
PS Claim 15; SEQ ID NO 8; 327pp; English.
XX
XX The present invention relates to a method (M1) for producing glucosamine
XX and N-acetylglucosamine by fermentation. The method comprises (a)
XX culturing in a fermentation medium a microorganism (I) which comprises
XX endogenous glucosamine-6-phosphate acetyltransferase (II) and a genetic
XX modification that increases the activity of (II), glucosamine-6-phosphate
XX synthase (III) or glucosamine-6-phosphate deaminase (IV), or decreases
XX the activity of (IV) and increases the activity of glucosamine-1
XX phosphate N-acetyltransferase (V), and (b) and collecting the product,
XX which is chosen from the group consisting of glucosamine-6-phosphate, N-
XX glucosamine, glucosamine-1-phosphate, N-acetylglucosamine-1-phosphate, N-
XX acetylglucosamine-6-phosphate, and N-acetylglucosamine. The present
XX sequence was used to illustrate the method of the invention.
SQ
Sequence 609 AA;

Query Match 99.8%; Score 3079; DB 8; Length 609;
Best Local Similarity 99.8%; Pred. No. 2e-263;
Matches 608; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MCGIVGAIAQDVAEILLEGRLRYGYDSAGLAVVDAEGHMTLRRLRGKVMQAQAAE 60
Db 1 MCGIVGAIAQDVAEILLEGRLRYGYDSAGLAVVDAEGHMTLRRLRGKVMQAQAAE 60

Qy 61 EPHLHGCGTGAHTRWATHGEPSEVNAHPHVSHEHVVVHNGIIEHHEPLREELKARGYTFV 120
Db 61 EPHLHGCGTGAHTRWATHGEPSEVNAHPHVSHEHVVVHNGIIEHHEPLREELKARGYTFV 120

| | | | |
|-----------------------------------|---------------------|--|-----|
| QY | 121 | SETDTEVIAHLVNNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPLV | 180 |
| Db | 121 | SETDTEVIAHLVNNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPLV | 180 |
| QY | 181 | IGLGMENFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIFDKTGAEVKRODIESNLQ | 240 |
| Db | 181 | IGLGMENFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIFDKTGAEVKRODIESNLQ | 240 |
| QY | 241 | YDAGDKGIYRHYWQKEIYEOPNAIKNTLTGRISHGQVDLSLSELPNADELLEKSKVEHIQILA | 300 |
| Db | 241 | YDAGDKGIYRHYWQKEIYEOPNAIKNTLTGRISHGQVDLSLSELPNADELLEKSKVEHIQILA | 300 |
| QY | 301 | CGTSYNSGMVSRVWFESLAGIPCDVEIASFRYKSAVRNSLMTLSQSGETADTTLA | 360 |
| Db | 301 | CGTSYNSGMVSRVWFESLAGIPCDVEIASFRYKSAVRNSLMTLSQSGETADTTLA | 360 |
| QY | 361 | RLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFPTTQLTVLLMLVAKL | 420 |
| Db | 361 | RLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFPTTQLTVLLMLVAKL | 420 |
| QY | 421 | SRLLKGLDASIEHDIVHGLQALPSRIEQMLSDQKRIEALAEFSDKXHALFLGRGDQYPIA | 480 |
| Db | 421 | SRLLKGLDASIEHDIVHGLQALPSRIEQMLSDQKRIEALAEFSDKXHALFLGRGDQYPIA | 480 |
| QY | 481 | LEGALKLKEISYTHAEAYAAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEEVRAR | 540 |
| Db | 481 | LEGALKLKEISYTHAEAYAAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEEVRAR | 540 |
| QY | 541 | GGQLYVFADQAGFVSSDNMHIEMPHVEEVIAPIFTYVPLQLLAYHVALIKGTDVQDPR | 600 |
| Db | 541 | GGQLYVFADQAGFVSSDNMHIEMPHVEEVIAPIFTYVPLQLLAYHVALIKGTDVQDPR | 600 |
| QY | 601 | NLAKSVTVE 609 | |
| Db | 601 | NLAKSVTVE 609 | |
| RESULT 14 | | | |
| AA5823 standard; protein; 609 AA. | | | |
| ID | AA5823 | | |
| XX | AC | | |
| XX | AA5823; | | |
| DT | 08-MAY-2000 | (first entry) | |
| XX | DE | E. coli glucosamine-6-phosphate synthase mutant Gln6P-S-49. | |
| XX | KW | Glucosamine-6-phosphate synthase; glnS gene; mutant; Gln6P-S-49; | |
| XX | KW | glucosamine; metabolic engineering; plasmid pKLN23-49; mutein. | |
| XX | OS | Escherichia coli. | |
| XX | Key | Location/Qualifiers | |
| FT | Misc-difference 4 | /note= "replaces wild-type Ile" | |
| FT | Misc-difference 272 | /note= "replaces wild-type 272" | |
| FT | Misc-difference 450 | /note= "replaces wild-type Ser" | |
| FT | WT0200004182-A1. | | |
| XX | PD | 27-JAN-2000. | |
| XX | PF | 15-JUL-1999; 99WO-US015976. | |
| XX | PR | 15-JUL-1998; 98US-00115475. | |
| XX | PA | (DCVB-) DCV INC DBA BIO-TECH RESOURCES. | |
| XX | PI | Berry A, Burlingame RP, Millis JR; | |
| XX | PI | | |

| | | | |
|-----------|---|---|-----|
| DR | WPI; 2000-182441/16. | | |
| DR | N-PSDB; AA258250. | | |
| XX | Fermentation of E. coli having an altered amino acid sugar metabolic pathway to produce glucosamine, especially using novel recombinant variant glucosamine-6-phosphate synthases. | | |
| PT | Claim 28; Page 117-119; 150pp; English. | | |
| XX | The present sequence is that of a mutant, denoted Gln6P-S-49, of the glucosamine-6-phosphate synthase (Gln6P synthase) of Escherichia coli. When compared with the wild-type sequence (see AA5822), the mutant includes 147, 1272T and 5450P amino acid substitutions. These alterations are predicted from the mutated glnS gene in plasmid pKLN23-49 (see AA58250). The invention provides methods for the overproduction of glucosamine by fermentation using a genetically engineered microorganism, especially E. coli, that includes a modified Gln6P synthase. Production of the glucosamine by recombinant strain 2123-49 (pKLN23-49) was significantly increased when compared to a strain expressing wild-type Gln6P synthase owing to reduced product inhibition | | |
| XX | Sequence 609 AA; | | |
| QY | Query Match | 99.5%; Score 3071; DB 3; Length 609; | |
| Db | Best Local Similarity | 99.5%; Pred. No. 1e-262; | |
| QY | Mismatches | 606; Conservative 0; Mismatches 3; Indels 0; Gaps 0; | |
| Db | 1 | MCICVIAIQRDVAEITLLEGLRELYGYDSAGLAVVDAEGHMTLRRLKGVQMLAQAAE | 60 |
| QY | 61 | EHPHGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHHEPRLKARGYTFV | 120 |
| Db | 61 | EHPHGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHHEPRLKARGYTFV | 120 |
| QY | 121 | SETDTEVIAHLVNNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPLV | 180 |
| Db | 121 | SETDTEVIAHLVNNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPLV | 180 |
| QY | 181 | IGLGMENFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIFDKTGAEVKRODIESNLQ | 240 |
| Db | 181 | IGLGMENFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIFDKTGAEVKRODIESNLQ | 240 |
| QY | 241 | YDAGDKGIYRHYWQKEIYEOPNAIKNTLTGRISHGQVDLSLSELPNADELLEKSKVEHIQILA | 300 |
| Db | 241 | YDAGDKGIYRHYWQKEIYEOPNAIKNTLTGRISHGQVDLSLSELPNADELLEKSKVEHIQILA | 300 |
| QY | 301 | CGTSYNSGMVSRVWFESLAGIPCDVEIASFRYKSAVRNSLMTLSQSGETADTTLA | 360 |
| Db | 301 | CGTSYNSGMVSRVWFESLAGIPCDVEIASFRYKSAVRNSLMTLSQSGETADTTLA | 360 |
| QY | 361 | RLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFPTTQLTVLLMLVAKL | 420 |
| Db | 361 | RLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFPTTQLTVLLMLVAKL | 420 |
| QY | 421 | SRLLKGLDASIEHDIVHGLQALPSRIEQMLSDQKRIEALAEFSDKXHALFLGRGDQYPIA | 480 |
| Db | 421 | SRLLKGLDASIEHDIVHGLQALPSRIEQMLSDQKRIEALAEFSDKXHALFLGRGDQYPIA | 480 |
| QY | 481 | LEGALKLKEISYTHAEAYAAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEEVRAR | 540 |
| Db | 481 | LEGALKLKEISYTHAEAYAAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEEVRAR | 540 |
| QY | 541 | GGQLYVFADQAGFVSSDNMHIEMPHVEEVIAPIFTYVPLQLLAYHVALIKGTDVQDPR | 600 |
| Db | 541 | GGQLYVFADQAGFVSSDNMHIEMPHVEEVIAPIFTYVPLQLLAYHVALIKGTDVQDPR | 600 |
| QY | 601 | NLAKSVTVE 609 | |
| Db | 601 | NLAKSVTVE 609 | |
| RESULT 15 | | | |

| | |
|--|---|
| ADI38859 | |
| ID | ADI38859 standard; protein; 609 AA. |
| XX | |
| AC | ADI38859; |
| XX | |
| DT | 15-APR-2004 (first entry) |
| XX | |
| DE | Mutant glucosamine-6-phosphate synthase, glms*49, SEQ ID 4. |
| XX | |
| KW | Glucosamine; N-acetylglucosamine; fermentation; |
| KW | glucosamine-6-phosphate acetyltransferase; |
| KW | glucosamine-6-phosphate synthase; glucosamine-6-phosphate deaminase; |
| KW | glucosamine-1-phosphate N-acetyltransferase; glucosamine-6-phosphate; |
| KW | glucosamine-1-phosphate; N-acetylglucosamine-1-phosphate; |
| XX | N-acetylglucosamine-6-phosphate; enzyme. |
| XX | |
| OS | Escherichia coli. |
| OS | Synthetic. |
| XX | |
| PN | WO2004003175-A2. |
| XX | |
| XX | |
| PD | 08-JAN-2004. |
| XX | |
| XX | |
| PF | 01-JUL-2003; 2003WO-US020925. |
| XX | |
| PR | 01-JUL-2002; 2002US-0393348P. |
| XX | |
| PA | (ARKI-) ARKION LIFE SCI LLC. |
| XX | |
| PI | Deng M, Angerer JD, Cyron D, Grund AD, Jerrell TA, Leanna C; |
| PI | Mathre O, Rossen R, Running J, Severson D, Song L, Wassink S; |
| XX | |
| DR | WPI; 2004-203380/19. |
| DR | N-PSDB; ADI38858. |
| XX | |
| PT | Producing glucosamine or N-acetylglucosamine by fermentation involves |
| PT | culturing microorganism comprising glucosamine-6-phosphate |
| PT | acetyltransferase, in fermentation medium, and collecting product. |
| XX | |
| PS | Claim 15; SEQ ID NO 4; 327pp; English. |
| XX | |
| CC | The present invention relates to a method (M1) for producing glucosamine |
| CC | and N-acetylglucosamine by fermentation. The method comprises (a) |
| CC | culturing in a fermentation medium a microorganism (I) which comprises |
| CC | endogenous glucosamine-6-phosphate acetyltransferase (II) and a genetic |
| CC | modification that increases the activity of (II), glucosamine-6-phosphate |
| CC | synthase (III) or glucosamine-6-phosphate deaminase (IV), or decreases |
| CC | the activity of (IV) and increases the activity of glucosamine-1 |
| CC | phosphate N-acetyltransferase (V), and (b) and collecting the product, |
| CC | which is chosen from the group consisting of glucosamine-6-phosphate, |
| CC | glucosamine, glucosamine-1-phosphate, N-acetylglucosamine-1-phosphate, N- |
| CC | acetylglucosamine-6-phosphate, and N-acetylglucosamine. The present |
| XX | sequence was used to illustrate the method of the invention. |
| SQ | Sequence 609 AA; |
| Query Match 99.5%; Score 3071; DB 8; Length 609; | |
| Best Local Similarity 99.5%; Pred. No. 1e-262; | |
| Matches 606; Conservative 0; Mismatches 3; Indels 0; Gaps 0; | |
| QY | 1 MCGIVGAIQRDVAEILLEGRLRLEYRGYDSAGLAVVDAEGHMTLRRLRGKVQMLAQAAE 60 |
| DB | 1 MCGTVGAIQRDVAEILLEGRLRLEYRGYDSAGLAVVDAEGHMTLRRLRGKVQMLAQAAE 60 |
| QY | 61 EHPHGGTGIAHTRWATHGEPSEVNAPHVSHVIVVHNGIIEENHEPLREELKARGYTFV 120 |
| DB | 61 EHPHGGTGIAHTRWATHGEPSEVNAPHVSHVIVVHNGIIEENHEPLREELKARGYTFV 120 |
| QY | 121 SETDTEVIAHLVNWELKQGGTREAIVRAIPOLRGAYGTIVIMDSRHPDPTLLAARSGSPLV 180 |
| DB | 121 SETDTEVIAHLVNWELKQGGTREAIVRAIPOLRGAYGTIVIMDSRHPDPTLLAARSGSPLV 180 |
| QY | 181 IGLGMGENFIASDQLALLPVTRRFIFLEBGDIAEITRRSVNIFDKTGAEVKQDIESNLQ 240 |

| | |
|----|--|
| DB | 181 IGLGMGENFIASDQLALLPVTRRFIFLEBGDIAEITRRSVNIFDKTGAEVKQDIESNLQ 240 |
| QY | 241 YDAGDKGIYRHYMQKEIYEOPNAIKNTLTGRIHQGVLDSELGPNADLLSKVEHIQILA 300 |
| DB | 241 YDAGDKGIYRHYMQKEIYEOPNAIKNTLTGRTSHQGVLDSELGPNADLLSKVEHIQILA 300 |
| QY | 301 CGTSYNSGMVSRWYFESLAGIPCDVEIASSEFRYKSAVRNLSMITLSQSGETADTLAGL 360 |
| DB | 301 CGTSYNSGMVSRWYFESLAGIPCDVEIASSEFRYKSAVRNLSMITLSQSGETADTLAGL 360 |
| QY | 361 RLSKELGYLGSLAICNVFGSSILVRESDLALMTNAGTEIGVASTKAFITQTLTVLLMLVAKL 420 |
| DB | 361 RLSKELGYLGSLAICNVFGSSILVRESDLALMTNAGTEIGVASTKAFITQTLTVLLMLVAKL 420 |
| QY | 421 SRLKGLDASIEHDIIVHGLQALPSRIEQMLPQDKRIEALAEFSDKHAFILGRGQOYPTA 480 |
| DB | 421 SRLKGLDASIEHDIIVHGLQALPSRIEQMLPQDKRIEALAEFSDKHAFILGRGQOYPTA 480 |
| QY | 481 LEGALKLKEISYIHAEEYAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEEVRAR 540 |
| DB | 481 LEGALKLKEISYIHAEEYAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEEVRAR 540 |
| QY | 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIPIFYTVPLQLLAYHVALIKGTVDQPR 600 |
| DB | 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIPIFYTVPLQLLAYHVALIKGTVDQPR 600 |
| QY | 601 NLAKSVTVE 609 |
| DB | 601 NLAKSVTVE 609 |

Search completed: August 7, 2006, 09:28:53
Job time : 96 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: August 7, 2006, 09:25:03 ; Search time 25 Seconds
(without alignments)
2343.840 Million cell updates/sec

Title: US-10-612-779-2

Perfect score: 3086

Sequence: 1 MCGIVGAIQAQRDVAEILLEG.....LIKGTDVDQPNLAKSVTVE 609

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 80.*

1: piri.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|--------|-------------|--------|----|--------|--------------------|
| 1 | 3086 | 100.0 | 609 | 1 | XNECGM | glutamine-fructose |
| 2 | 3082 | 99.9 | 609 | 2 | H86058 | hypothetical prote |
| 3 | 3075 | 99.6 | 609 | 2 | G91212 | hypothetical prote |
| 4 | 3053 | 98.9 | 609 | 2 | AB0955 | glutamine-fructose |
| 5 | 2682 | 86.9 | 609 | 2 | AB0500 | glutamine-fructose |
| 6 | 2270.5 | 73.6 | 610 | 2 | D4067 | glutamine-fructose |
| 7 | 2264.5 | 73.4 | 610 | 2 | B82316 | glutamine-fructose |
| 8 | 1995 | 64.6 | 611 | 2 | F82951 | glutamine-fructose |
| 9 | 1868 | 60.5 | 621 | 2 | A84933 | glutamine-fructose |
| 10 | 1702 | 55.2 | 635 | 2 | E71272 | glutamine-fructose |
| 11 | 1695 | 54.9 | 609 | 2 | A82844 | glutamine-fructose |
| 12 | 1597 | 51.7 | 611 | 2 | T45493 | glutamine-fructose |
| 13 | 1570.5 | 50.9 | 612 | 2 | H82022 | glutamine-fructose |
| 14 | 1566.5 | 50.8 | 612 | 2 | B81246 | glutamine-fructose |
| 15 | 1449 | 47.0 | 612 | 2 | AD3595 | glutamine-fructose |
| 16 | 1435.5 | 46.3 | 608 | 2 | B95322 | NodM Glutamine ami |
| 17 | 1427.5 | 46.3 | 608 | 2 | C97575 | glutamine-fructose |
| 18 | 1427.5 | 46.3 | 608 | 2 | AC2796 | hypothetical prote |
| 19 | 1411.5 | 45.7 | 604 | 2 | G69793 | glutamine-fructose |
| 20 | 1371.5 | 44.4 | 606 | 2 | A75336 | glutamine-fructose |
| 21 | 1361.5 | 44.1 | 592 | 2 | D70327 | glutamine-fructose |
| 22 | 1340.5 | 43.4 | 608 | 1 | S01040 | glutamine-fructose |
| 23 | 1339 | 43.4 | 605 | 2 | S16561 | glutamine-fructose |
| 24 | 1337.5 | 43.3 | 606 | 2 | B87263 | hypothetical prote |
| 25 | 1327.5 | 43.0 | 598 | 2 | F75212 | glutamine-fructose |
| 26 | 1314.5 | 42.6 | 606 | 2 | B72412 | glutamine-fructose |
| 27 | 1297.5 | 42.0 | 608 | 2 | C96919 | glutamine-fructose |
| 28 | 1293 | 41.9 | 615 | 2 | T35569 | glutamine-fructose |
| 29 | 1292.5 | 41.9 | 601 | 2 | D71248 | glutamine-fructose |

RESULT 1

XNECGM

glutamine-fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) - Escherichia coli
N;Alternate names: glucosamine-6-phosphate synthase; glucosaminophosphate isomerase (glu
e
C;Species: Escherichia coli
C;Date: 31-Mar-1990 #sequence_revision 17-Oct-1997 #text_change 09-Jul-2004
R;Accession: B65176; A30389; I41219; S17839; Q90513
R;Plattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: B65176
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-609 <BLAT>
A;Cross-references: UNIPROT:P17169; UNIPARC:UPI0000153CA3; GB:AE000450; GB:U00096; NID:g
A;Experimental source: strain K-12, substrain MG1655
R;Walker, J.E.; Gay, N.J.; Saraste, M.; Eberle, A.N.
Biochem. J. 224, 799-815, 1984
A;Title: DNA sequence around the Escherichia coli unc operon. Completion of the sequence
A;Reference number: A30389; MUID:85121806; PMID:6395859
A;Accession: A30389
A;Molecule type: DNA
A;Residues: 1-418,'NV',421-609 <WAL>
A;Cross-references: UNIPARC:UPI000016F5F4; GB:X01631; NID:g43256; PIDN:CAA25785.1; PID:g
R;McKown, R.L.; Orle, K.A.; Chen, T.; Craig, N.L.
J. Bacteriol. 170, 352-358, 1988
A;Title: Sequence requirements of Escherichia coli attTn7, a specific site of transposon
A;Reference number: I41219; MUID:88086894; PMID:2826397
A;Accession: I41219
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 597-609 <RES>
A;Cross-references: UNIPARC:UPI000016F18D; GB:M18980; NID:g146074; PIDN:AAA23836.1; PID:g
R;Golinelli-Pimpaneau, B.; Badet, B.
Eur. J. Biochem. 201, 175-182, 1991
A;Title: Possible involvement of Lys603 from Escherichia coli glucosamine-6-phosphate syr
A;Reference number: S17839; MUID:92007872; PMID:1915361
A;Accession: S17839
A;Molecule type: protein
A;Residues: 49-50,'X',52-53,218,219,220-223,'X',225-231,489-493,504-508,'K','K','K',601-609
4A7
A;Cross-references: UNIPARC:UPI000011EBBE; UNIPARC:UPI000011ECBE; UNIPARC:UPI000011ECBE;
C;Genetics:
A;Gene: gims
A;Map position: 4 min
C;Function:
A;Description: catalyzes the formation of D-glucosamine-6-phosphate from the amido group
A;Pathway: Glucosamine biosynthesis
A;Note: Glucosamine-6-phosphate is used in the biosynthesis of amino sugars of asparagine

| | |
|--|---|
| C;Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing) | |
| C;Keywords: aminotransferase; isomerase | |
| F;2-609/Product: glutamine-fructose-6-phosphate transaminase (isomerizing) #status predicted | |
| F;2/Active site: Cys #status predicted | |
| Query Match | 100.0%; Score 3086; DB 1; Length 609; |
| Best Local Similarity | 100.0%; Pred. No. 2e-191; |
| Matches 609; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | |
| Qy | 1 MCGIVGAIAQRDVAEILLEGRLREYRGYDSAGLAVVDAEGHMTLRRLRGKQVQMLAQAAE 60 |
| Db | 1 MCGIVGAIAQRDVAEILLEGRLREYRGYDSAGLAVVDAEGHMTLRRLRGKQVQMLAQAAE 60 |
| Qy | 61 EHPHGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHHEPLREELKARGYTFV 120 |
| Db | 61 EHPHGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHHEPLREELKARGYTFV 120 |
| Qy | 121 SETDTEVIAHLVNWELKQGGLTREAIVRAIPQLRGAYGTVMDSRHPDTLAARSGSPLV 180 |
| Db | 121 SETDTEVIAHLVNWELKQGGLTREAIVRAIPQLRGAYGTVMDSRHPDTLAARSGSPLV 180 |
| Qy | 181 IGLGMENFTASQDALLPVTRRFIFLEEGDIAEITRRSVNIPDKTGAEVKQDIESNLQ 240 |
| Db | 181 IGLGMENFTASQDALLPVTRRFIFLEEGDIAEITRRSVNIPDKTGAEVKQDIESNLQ 240 |
| Qy | 241 YDAGDKGIYRHYMQKEIYEQPNAIKNTLTGRISHGOVDLSELGPNADLLSKVEHIQILA 300 |
| Db | 241 YDAGDKGIYRHYMQKEIYEQPNAIKNTLTGRISHGOVDLSELGPNADLLSKVEHIQILA 300 |
| Qy | 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASSEFRYKSAVRNSLMITLSQSGETADTLA 360 |
| Db | 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASSEFRYKSAVRNSLMITLSQSGETADTLA 360 |
| Qy | 361 RLSKELGYLGS LAICNVPGSSLVRESDLALMTNAGTEIGVASTKATFTQLTVLLMLVAKL 420 |
| Db | 361 RLSKELGYLGS LAICNVPGSSLVRESDLALMTNAGTEIGVASTKATFTQLTVLLMLVAKL 420 |
| Qy | 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDOKRIEALAEFSDKHAFFLGRGQDYP 480 |
| Db | 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDOKRIEALAEFSDKHAFFLGRGQDYP 480 |
| Qy | 481 LEGALKKEISYTHABAYAAAGELKHGPLALIDAMPVIVVAPNNELLEKLSNIEVRAR 540 |
| Db | 481 LEGALKKEISYTHABAYAAAGELKHGPLALIDAMPVIVVAPNNELLEKLSNIEVRAR 540 |
| Qy | 541 GGQLYVFADQDAGFVSSDNMHIIEMPHVEEVIAPFYFTVPLQLLAYHVALIKGTDVDQPR 600 |
| Db | 541 GGQLYVFADQDAGFVSSDNMHIIEMPHVEEVIAPFYFTVPLQLLAYHVALIKGTDVDQPR 600 |
| Qy | 601 NLAKSVTVE 609 |
| Db | 601 NLAKSVTVE 609 |
| RESULT 2 | |
| H86058 | |
| hypothetical protein glms [imported] - Escherichia coli (strain O157:H7, substrain EDL93 | |
| C;Species: Escherichia coli | |
| C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004 | |
| C;Accession: H86058 | |
| R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew | |
| iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, | |
| Nature 409, 529-533, 2001 | |
| A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7. | |
| A;Reference number: A85480; MUID:21074935; PMID:11206551 | |
| A;Accession: H86058 | |
| A;Status: preliminary | |
| A;Molecule type: DNA | |
| A;Residues: 1-609 <STO> | |
| A;Cross-references: UNIPROT:Q8XEG2; UNIPARC:UPI0000165991; GB:AE005174; NID:g12518583; F | |
| C;Experimental source: strain O157:H7, substrain EDL933 | |
| C;Genetics: | |
| A;Gene: glms | |

| | |
|--|---|
| C;Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing) | |
| C;Keywords: aminotransferase; isomerase | |
| F;2-609/Product: glutamine-fructose-6-phosphate transaminase (isomerizing) #status predicted | |
| F;2/Active site: Cys #status predicted | |
| Query Match | 99.9%; Score 3082; DB 2; Length 609; |
| Best Local Similarity | 99.8%; Pred. No. 3.7e-191; |
| Matches 608; Conservative 1; Mismatches 0; Indels 0; Gaps 0; | |
| Qy | 1 MCGIVGAIAQRDVAEILLEGRLREYRGYDSAGLAVVDAEGHMTLRRLRGKQVQMLAQAAE 60 |
| Db | 1 MCGIVGAIAQRDVAEILLEGRLREYRGYDSAGLAVVDAEGHMTLRRLRGKQVQMLAQAAE 60 |
| Qy | 61 EHPHGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHHEPLREELKARGYTFV 120 |
| Db | 61 EHPHGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHHEPLREELKARGYTFV 120 |
| Qy | 121 SETDTEVIAHLVNWELKQGGLTREAIVRAIPQLRGAYGTVMDSRHPDTLAARSGSPLV 180 |
| Db | 121 SETDTEVIAHLVNWELKQGGLTREAIVRAIPQLRGAYGTVMDSRHPDTLAARSGSPLV 180 |
| Qy | 181 IGLGMENFTASQDALLPVTRRFIFLEEGDIAEITRRSVNIPDKTGAEVKQDIESNLQ 240 |
| Db | 181 IGLGMENFTASQDALLPVTRRFIFLEEGDIAEITRRSVNIPDKTGAEVKQDIESNLQ 240 |
| Qy | 241 YDAGDKGIYRHYMQKEIYEQPNAIKNTLTGRISHGOVDLSELGPNADLLSKVEHIQILA 300 |
| Db | 241 YDAGDKGIYRHYMQKEIYEQPNAIKNTLTGRISHGOVDLSELGPNADLLSKVEHIQILA 300 |
| Qy | 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASSEFRYKSAVRNSLMITLSQSGETADTLA 360 |
| Db | 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASSEFRYKSAVRNSLMITLSQSGETADTLA 360 |
| Qy | 361 RLSKELGYLGS LAICNVPGSSLVRESDLALMTNAGTEIGVASTKATFTQLTVLLMLVAKL 420 |
| Db | 361 RLSKELGYLGS LAICNVPGSSLVRESDLALMTNAGTEIGVASTKATFTQLTVLLMLVAKL 420 |
| Qy | 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDOKRIEALAEFSDKHAFFLGRGQDYP 480 |
| Db | 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDOKRIEALAEFSDKHAFFLGRGQDYP 480 |
| Qy | 481 LEGALKKEISYTHABAYAAAGELKHGPLALIDAMPVIVVAPNNELLEKLSNIEVRAR 540 |
| Db | 481 LEGALKKEISYTHABAYAAAGELKHGPLALIDAMPVIVVAPNNELLEKLSNIEVRAR 540 |
| Qy | 541 GGQLYVFADQDAGFVSSDNMHIIEMPHVEEVIAPFYFTVPLQLLAYHVALIKGTDVDQPR 600 |
| Db | 541 GGQLYVFADQDAGFVSSDNMHIIEMPHVEEVIAPFYFTVPLQLLAYHVALIKGTDVDQPR 600 |
| Qy | 601 NLAKSVTVE 609 |
| Db | 601 NLAKSVTVE 609 |
| RESULT 3 | |
| G91212 | |
| hypothetical protein Ecs4671 [imported] - Escherichia coli (strain O157:H7, substrain RIN | |
| C;Species: Escherichia coli | |
| C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004 | |
| C;Accession: G91212 | |
| R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; | |
| gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H. | |
| DNA Res. 8, 11-22, 2001 | |
| A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genom | |
| A;Reference number: A99629; MUID:21156231; PMID:11258796 | |
| A;Accession: G91212 | |
| A;Status: preliminary | |
| A;Molecule type: DNA | |
| A;Residues: 1-609 <HAY> | |
| A;Cross-references: UNIPROT:Q8XEG2; UNIPARC:UPI000016552F; GB:BA000007; PIDN:BA838094.1; | |
| A;Experimental source: strain O157:H7, substrain RIMD 0509952 | |
| C;Genetics: | |
| A;Gene: Ecs4671 | |
| C;Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing) | |
| Query Match | 99.6%; Score 3075; DB 2; Length 609; |

| | Best Local Similarity | 99.7%; | Pred. No. le-190; | Matches | 607; | Conservative | 2; | Mismatches | 0; | Indels | 0; | Gaps | 0; | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|----|-----------------------|--------|-------------------|---------|------|--------------|----|------------|-----|--------|-----|------|----|-----|----|---|----|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|----|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|-----|-----|-----|---|---|-----|-----|
| Qy | 1 | MC | GIVGAIAQ | RDV | AEI | LL | EG | LR | LEY | RG | YDS | AG | LA | VVD | AE | B | CH | M | T | R | R | L | G | K | V | Q | M | L | A | A | A | E | 60 | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Db | 1 | MC | GIVGAIAQ | RDV | AEI | LL | EG | LR | LEY | RG | YDS | AG | LA | VVD | AE | B | CH | M | T | R | R | L | G | K | V | Q | M | L | A | A | A | E | 60 | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Qy | 61 | E | PH | L | H | G | G | T | G | I | A | H | T | W | A | T | H | G | B | E | P | S | V | N | A | H | P | H | S | E | H | I | V | V | H | N | G | I | I | E | N | H | E | P | L | R | E | L | K | A | R | G | T | T | V | 120 | | | | | |
| Db | 61 | E | PH | L | H | G | G | T | G | I | A | H | T | W | A | T | H | G | B | E | P | S | V | N | A | H | P | H | S | E | H | I | V | V | H | N | G | I | I | E | N | H | E | P | L | R | E | L | K | A | R | G | T | T | V | 120 | | | | | |
| Qy | 121 | S | E | T | D | E | V | I | A | H | L | V | N | N | E | L | K | O | G | G | T | R | E | A | V | L | R | A | T | P | O | L | E | G | A | G | T | V | I | M | D | S | R | H | P | D | T | T | L | A | R | S | S | P | L | V | 180 | | | | |
| Db | 121 | S | E | T | D | E | V | I | A | H | L | V | N | N | E | L | K | O | G | G | T | R | E | A | V | L | R | A | T | P | O | L | E | G | A | G | T | V | I | M | D | S | R | H | P | D | T | T | L | A | R | S | S | P | L | V | 180 | | | | |
| Qy | 181 | I | G | L | G | M | G | N | F | I | A | S | D | O | L | A | L | P | V | T | R | R | F | I | F | L | E | E | G | D | I | A | E | I | T | R | R | S | V | N | I | F | D | K | T | G | A | E | V | K | R | O | D | I | E | S | N | L | Q | 240 | |
| Db | 181 | I | G | L | G | M | G | N | F | I | A | S | D | O | L | A | L | P | V | T | R | R | F | I | F | L | E | E | G | D | I | A | E | I | T | R | R | S | V | N | I | F | D | K | T | G | A | E | V | K | R | O | D | I | E | S | N | L | Q | 240 | |
| Qy | 241 | Y | D | A | D | K | G | I | G | I | R | H | M | O | K | E | I | Y | E | O | P | N | A | I | K | N | T | L | T | G | R | I | S | H | G | Q | V | D | L | S | E | L | G | P | N | A | D | E | L | L | S | K | V | E | H | I | Q | I | L | A | 300 |
| Db | 241 | Y | D | A | D | K | G | I | G | I | R | H | M | O | K | E | I | Y | E | O | P | N | A | I | K | N | T | L | T | G | R | I | S | H | G | Q | V | D | L | S | E | L | G | P | N | A | D | E | L | L | S | K | V | E | H | I | Q | I | L | A | 300 |
| Qy | 301 | C | G | T | S | Y | N | G | M | S | R | Y | F | E | S | I | A | G | I | P | C | D | V | E | T | A | S | S | F | R | K | S | A | V | R | N | S | L | M | I | T | L | S | O | S | E | T | A | D | T | L | A | G | L | 360 | | | | | | |
| Db | 301 | C | G | T | S | Y | N | G | M | S | R | Y | F | E | S | I | A | G | I | P | C | D | V | E | T | A | S | S | F | R | K | S | A | V | R | N | S | L | M | I | T | L | S | O | S | E | T | A | D | T | L | A | G | L | 360 | | | | | | |
| Qy | 361 | R | L | S | K | E | L | G | L | G | S | L | A | I | C | N | P | G | S | S | L | R | E | S | D | L | A | M | T | N | A | G | T | E | I | G | V | A | S | T | K | A | F | T | T | O | L | T | V | L | L | M | V | A | K | L | 420 | | | | |
| Db | 361 | R | L | S | K | E | L | G | L | G | S | L | A | I | C | N | P | G | S | S | L | R | E | S | D | L | A | M | T | N | A | G | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 4
AB0955
glutamine-fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) - Salmonella ent
N:Alternate names: Glucosamine-fructose-6-P aminotransferase [misnomer]
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 28-Jul-2003
C:Accession: AB0955
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AB0955
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-609 <PAR>
A:Cross-references: UNIPARC:UPI000005A683; GB:AL513382; PIDN:CAD03134.1; PID:G16504770;
C:Genetics:
A:Gene: STY3917
C:Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing)
C:Keywords: aminotransferase; isomerase

| Query Match | 98.9% | Score 3053; | DB 2; | Length 609; |
|-----------------------|-----------------|-----------------------|--|-------------|
| Best Local Similarity | 98.5%; | Pred. No. 2.7e-189; | | |
| Matches 600; | Conservative 5; | Mismatches 4; | Indels 0; | Gaps 0; |
| QY | 1 | MCGIVGAIAORDVAEII | LEGLRRLEYRGYDSAGLAVDAEGHMTLRRLRGKVQMLAAQAE | 60 |
| DB | 1 | MCGIVGAIAORDVAEII | LEGLRRLEYRGYDSAGLAVDAEGHMTLRRLRGKVQMLAAQAE | 60 |
| QY | 61 | EHPHGGGTGIAHTRWATHG | PPSEVNAHPHVESEHIVVHNGIIENHEPLREBELKARGTTFV | 120 |
| DB | 61 | EHPHGGGTGIAHTRWATHG | PPSEANAHPHVESEHIVVHNGIIENHEPLREALKARGTTFV | 120 |
| QY | 121 | SETDETVIAHLVNVELKQGGT | FLREAVLRAIPQLRGAYGTVMDSRHPDPTLLAARSGSPLV | 180 |
| DB | 121 | SETDETVIAHLVNVELKQGGT | FLRDIAIIRAIPQLRGAYGTVMDRHPDPTLLAARSGSPLV | 180 |
| QY | 181 | IGLWGNFIASDQALALPVT | RRFIFLEEGDIABITRRSVNIFDKTGAEVKRDIESNLQ | 240 |
| DB | 181 | IGLWGNFIASDQALALPVT | RRFIFLEEGDIABITRRSVNIFDNTGAEVKRDIESNLQ | 240 |
| QY | 241 | YDAGDKGIYRHYMOKE | IYEOPNAIKNTLGRI SHGOVDLSLGNADLLSKVHEHIOILA | 300 |
| DB | 241 | YDAGDKGIYRHYMOKE | IYEOPNAIKNTLGRI SHGOVDLSLGNADLLSKVHEHIOILA | 300 |
| QY | 301 | CGTSYNSGWSRYWFESLAGI | PCDVEIASEFRYKSAVRNSLMTILTSQGETADTLTLAGL | 360 |
| DB | 301 | CGTSYNSGWSRYWFESLAGI | PCDVEIASEFRYKSAVRNSLMTILTSQGETADTLTLAGL | 360 |
| QY | 361 | RLSKELGYLSLAI | CNVPGSSLVRESDLALMTNAGTEIGVASTKAFITOLTVLIMLVAKL | 420 |
| DB | 361 | RLSKELGYLSLAI | CNVPGSSLVRESDLALMTNAGTEIGVASTKAFITOLTVLIMLVAKL | 420 |
| QY | 421 | SRLKGLDASIEHDIVHGLQAL | PSRIEOMLSODKRIEALAEPDFSKHHALFLGRGDQVPIA | 480 |
| DB | 421 | SRLKGLDASIEHDIVHGLQAL | PSRIEOMLSODKRIEALAEPDFSKHHALFLGRGDQVPIA | 480 |
| QY | 481 | LEGALKUKETSITHAEVYAA | GELKHGPALIDADMPVIVVAPNNLELEKLSNTEEVRRAR | 540 |
| DB | 481 | LEGALKUKETSITHAEVYAA | GELKHGPALIDADMPVIVVAPNNLELEKLSNTEEVRRAR | 540 |
| QY | 541 | GGQYLVFADODDAGFVSSD | NMHIIEMPHVEEVIAPIFVTVPQLLAYHVALLKGTDVDQPR | 600 |
| DB | 541 | GGQYLVFSDQDAGFVSSD | NMHIIEMPHVEEVIAPIFVTVPQLLAYHVALLKGTDVDQPR | 600 |
| QY | 601 | NLAKSVTVE | 609 | |
| DB | 601 | NLAKSVTVE | 609 | |

```

RESULT 5
AB0500
glutamine-fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) [import
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C:Accession: AB0500
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentis-
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dou-
ll, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.;
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; UID:21470413; PMID:11586360
A:Accession: AB0500
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-609 <KUR>
A:Cross-references: UNIPROT:Q8Z9S8; UNIPARC:UPI0000165B7A; GB:AL590842; PIDN:CA
C:Genetics:
C:Gene: glms
C:Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing)
C:Keywords: aminotransferase; isomerase

Query Match      86.9%  Score 2682;  DB 2;  Length 609;
Best Local Similarity  85.1%  Pred. No. 2,4e-165;

```

| Matches | 518; | Conservative | 44; | Mismatches | 47; | Indels | 0; | Gaps | 0; | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---------|------|--------------|-----|------------|-----|--------|----|------|----|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|-----|----|---|-----|-----|-----|---|-----|-----|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|
| QY | 1 | MC | G | I | G | A | I | A | Q | R | D | V | A | E | I | L | L | E | G | L | R | L | E | Y | G | D | S | A | G | L | A | V | D | A | E | G | H | M | T | R | L | R | L | R | L | G | K | V | M | L | A | Q | A | E | 60 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| DB | 1 | MC | G | I | G | V | A | Q | R | D | V | A | E | I | L | L | E | G | L | R | L | E | Y | G | D | S | A | G | L | A | V | D | S | E | G | H | L | T | R | L | R | V | G | K | V | H | A | L | S | D | A | E | 60 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| QY | 61 | E | H | P | L | H | G | G | T | G | I | A | T | R | W | A | T | H | E | P | S | E | V | N | A | H | P | V | S | H | I | V | V | H | N | G | I | I | E | N | H | E | P | L | R | E | E | K | A | R | G | T | P | V | 120 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| DB | 61 | K | O | D | L | H | G | G | T | G | I | A | T | R | W | A | T | H | E | P | S | E | V | N | A | H | P | V | S | I | S | V | V | H | N | G | I | I | E | N | H | E | P | L | R | E | L | L | S | R | G | T | P | V | 120 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| QY | 121 | S | E | T | D | T | E | V | I | A | H | L | V | N | W | E | L | K | O | G | G | T | L | E | A | V | L | R | A | I | P | O | L | R | G | A | Y | G | T | V | I | M | D | S | R | H | P | T | L | L | A | A | R | S | G | S | P | L | V | 180 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| DB | 121 | S | E | T | D | T | E | V | I | A | H | L | V | N | W | E | Q | O | G | G | S | L | E | V | L | V | K | R | V | I | P | O | L | R | G | A | Y | G | T | V | I | M | D | S | R | H | P | T | L | L | A | A | R | S | G | S | P | L | V | 180 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| QY | 181 | I | G | L | G | E | N | F | I | A | S | D | O | L | L | P | V | T | R | R | F | I | L | E | G | D | I | A | E | T | R | R | S | V | N | I | F | D | T | K | G | A | E | V | R | K | O | D | I | E | S | N | L | Q | 240 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| DB | 181 | I | G | C | G | V | E | N | F | I | A | S | D | O | L | L | P | V | T | R | R | F | I | L | E | G | D | V | E | T | R | R | S | I | S | I | F | D | K | O | G | N | A | I | E | R | P | E | S | Q | V | 240 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| QY | 241 | Y | D | A | G | D | K | G | I | Y | R | H | Y | M | O | K | E | I | Y | E | O | P | N | A | I | K | N | T | L | T | G | R | I | S | H | G | O | V | D | L | S | E | L | G | P | N | A | D | E | L | L | 300 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| DB | 241 | Y | D | A | G | D | K | G | I | Y | R | H | Y | M | O | K | E | I | Y | E | O | P | N | A | I | K | N | T | L | E | G | R | I | S | H | G | M | I | D | L | S | E | L | G | P | K | A | D | A | L | L | A | E | V | H | I | 300 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| QY | 301 | C | G | T | S | N | G | M | V | S | R | Y | P | E | S | L | A | G | I | P | C | D | V | E | I | A | S | E | F | R | Y | K | S | A | V | R | R | N | S | L | M | I | T | L | S | O | S | E | T | A | D | L | A | G | L | 360 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| DB | 301 | C | G | T | S | N | G | M | V | S | R | Y | P | E | S | L | A | G | I | P | C | D | V | E | I | A | S | E | F | R | Y | K | S | A | V | R | P | N | S | L | L | I | T | L | S | O | S | E | T | A | D | L | A | L | 360 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| QY | 361 | R | L | S | K | E | L | G | V | L | G | S | L | A | I | C | N | V | P | G | S | S | I | V | R | E | S | D | L | A | M | T | N | A | G | T | E | I | G | V | A | S | T | K | A | F | T | T | O | L | V | I | L | M | V | A | K | L | 420 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| DB | 361 | R | L | S | K | E | L | G | V | L | G | S | L | A | I | C | N | V | A | G | S | S | L | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Matches 441; Conservative 73; Mismatches 95; Indels 1; Gaps 1;

QY 1 MCGIIVGAIAQRDVAETILLGLARLEVRGYDSAGLAVVDAEGHMTLRRLRKGVOMLAQAAE 60
Db | | | | | | | | | | | | | | | | : : : | | | | | : |
1 MCGIIVGAIAQRDVAETILLGLARLEVRGYDSAGLAVVDAEGHMTLRRLRKGVOMLAQAAE 60
QY 61 EHPLHGCTGIAHTRWATHGPSEVNAPHSVEBHVHHNGIIENHEPRLREELKARGTFFV 120
Db | | | | | | | | | | | | | | | | : | | | | | | | | | | : |
61 EKPLHGCTGIAHTRWATHGPSETNAHPSSGTFAVVHNGIIEHNEELRELLKSRGVVFL 120
QY 121 SETDTFVTAHLVNWELKQGCTIREAIVLRAIPOLRGAYGTVMDSRPHDPTLLAARSGLPV 180
Db | | | | | | | | | | : : : | | | | | : | | | | | : |
121 SOTDTFVTAHLVNEWEMRTDSLDDAKKAVKQLTGAYGMVMDSRHPHLVAARSGSPVL 180
QY 181 ICLMGGENFIASDQLALLPVTRRFIFLEGSDTAETRRSVNIFDKTGAEVKRKDIESNLQ 240
Db | | | | | | | | | | | | | | | | : | | | | | : | | | | : |
181 ICLGIGENFLASDQLALLSVTRRFIFLEGSDTAETRRVDIYDTGNKAKRIHESNLE 240
QY 241 YDAGDKGIYRHVMQKEIYEOPNAIKNTLTGRISHGOVDLSLGPNADPELLSKVEHIQILA 300
Db | | | | | | | | | | | | | | | | : | | | | | : | | | | : |
241 NDAAEKGFRHFMOKEIYEQPALTINTWEGRIHNHENVIDSICNGAKGILEKVEHIQIVA 300
QY 301 CGTSNSGWRSRYWFESLAGIPCDEVIASEPRYSKSAVERNSIMITLSOSGETADTLA 360
Db | | | | | | | | | | | | | | | | : | | | | | : | | | | : |
301 CGTSNAGWARYWFESLAGVSCDEVIASEFRYKFTVPNSLIITLSOSGETADTLAAL 360
QY 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFTTQLTVLLMLVAKL 420
Db | | | | | | | | | | | | | | | | : | | | | | : | | | | : |
361 RLAKKGYMAALTCINVAGSSLVRESDLAFMTRAGVEGVASTKAFTTQLAALLMLVTAL 420
QY 421 SRLKG-LDASTEHDIHVHGLQALPSRIEQMLSODKEIALAEDFSKHGHAFILGRGDQYPI 479
Db | | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
421 GKVGHHSIVEKERETLIKAWQLPAIEIKALAPDTIEALAEDFAEKHHAFILGRGAFYPI 480
QY 480 ALEGALKLKEISYTHAEAYAAGELKHGPLALIDAMPVIVVAPNNELLEKLKSNIIEEVR 539
Db | | | | | | | | | | | | | | | | : | | | | | : | | | | : |
481 AVEASLKLEISYTHAEAYAAGELKHGPLALIDAMPVIVVAPNNELLEKVKSNIIEEVR 540
QY 540 RGGOLYFPADODAGFVSDDNHIIEMPHVEVIAPIFTVPLOLLAYHVALIKGTDVDQP 599
Db | | | | | | | | | | | | | | | | : | | | | | : | | | | : |
541 RGGOLYFPADKEAGFTPEGMKIITMPKNDIVAPIFYTIQMQLLSYVALLIKGTDVDQP 600
QY 600 RNLAHSVTFVE 609
Db | | | | | | | | | | | | | | | | : | | | | | : | | | | : |
601 RNLAHSVTFVE 610

RESULT 7
E82316
glutamine-fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) [similarity] - Vj
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: E82316
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.
Nature 406, 477-483, 2000
A>Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: E82316
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-610 <HEI>
A:Cross-references: UNIPROT:Q9KUM6; UNIPARC:UIP0000164B53; GB:AE004135; GB:AE003852; NID:
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC0487
A:Map position: 1
C:Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing)
C:Keywords: aminotransferase; isomerase
F:2-610/Product: glutamine-fructose-6-phosphate transaminase (isomerizing) #status predic
F:2/Active site: Cys #status predicted

Query Match 73.4%; Score 2264.5; DB 2; Length 610;
Best Local Similarity 71.6%; Pred. No. 2.2e-138;
Matches 437; Conservative 72; Mismatches 100; Indels 1; Gaps 1;

Qy 1 MCGIVGIAQORDVAEILLEGRLREYRGYSAGLAVVDAEGHMTLRRLKGVQMLAAAE 60
Db 1 MCGIVGAVARDVAEILVQGLRLREYRGYSAGVAVVDSKQTLRLRLKGVQELADAVE 60

Qy 61 EHPHGGTGAHTRWATHGEPSEVNAHPVSHIIVVHNGIIEHPELREELKARGYTFV 120
Db 61 AAQVAGGTGAHTRWATHGEPSEINAPHISGDIIVVHNGIIEHPELRLTMDQGVYFT 120

Qy 121 SETDTEVIAHLVNWELKOGGTLREAVLRAIPOLRGAYGTVMDSRHPDPTLLAARSGPLV 180
Db 121 SQTDETEVIAHLVNWELKOGGTLREAVLRAIPOLRGAYGTVMDSRHPDPTLLAARSGPLV 180

Qy 181 IGLGMENFTASDQALLPVTFRFLEEGDIAETRRSVNIYFDKTAEGVKRODIESNLQ 240
Db 181 IGFIGENFTASDQALLPVTFRFLEEGDIAETRRSVNIYFDKTAEGVKRODIESNLQ 240

Qy 241 YDAGDKGIYRHYMQEIEYQPNNAIKNTLTGRISHGQVDSLSELGPNADLLSKVEHIQILA 300
Db 241 HDADKGRHYRHYMQEIEYQPNNAIKNTLTGRISHGQVDSLSELGPNADLLSKVEHIQILA 300

Qy 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASBFRYKSAVRNLSMITLSQSGETADTLAAGL 360
Db 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASBFRYKSAVRNLSMITLSQSGETADTLAAGL 360

Qy 361 RLSEKELGYLGLSLAICNVPGSSLVRESDLAMTNAGTEIGVASTKAFPTTQLTVLLMLVAKL 420
Db 361 RLSEKELGYLGLSLAICNVPGSSLVRESDLAMTNAGTEIGVASTKAFPTTQLTVLLMLVAKL 420

Qy 421 SR-LKGLDASIEHDIVHGLQALPSRIEOMLSODKRIEALAEFSDKHAFPLGRGQOYPI 479
Db 421 GKQOIRIGLEAEIIVHGLQALPSRIEOMLSODKRIEALAEFSDKHAFPLGRGQOYPI 479

Qy 480 ALEGALKKEISYTHAEYAAAGELKHGPLALIDADMPIVIVVAPNNELLEKLKSNIEEVR 539
Db 481 AVEASLKKEISYTHAEYAAAGELKHGPLALIDADMPIVIVVAPNNELLEKLKSNIEEVR 540

Qy 540 RGQQLYVAFADQAGFVSSDNMHIEMPHVEEVIAPITVPLQLLAYHVALIKGTVDQOP 599
Db 541 RGQQLYVAFADQAGFVSSDNMHIEMPHVEEVIAPITVPLQLLAYHVALIKGTVDQOP 600

Qy 600 RNLAKSVTVE 609
Db 601 RNLAKSVTVE 610

RESULT 8
F82951
Glutamine-fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) [similarity] - B
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: F82951
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Berman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen
A;Reference number: A82950; PMID:10384043
A;Accession: F82951
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-611 <STO>
A;Cross-references: UNIPROT:Q9HT25; UNIPARC:UPI0000165E0; GB:AE004091; NCBI:U00001.1
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: glms; PA5549
C;Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing)
C;Keywords: aminotransferase; isomerase
F;2-611/Product: glutamine-fructose-6-phosphate transaminase (isomerizing) #status predicted
F;2/Active site: Cys #status predicted

Query Match 64.6%; Score 1995; DB 2; Length 611;
Best Local Similarity 62.7%; Pred. No. 5.4e-121;
Matches 383; Conservative 101; Mismatches 125; Indels 2; Gaps 2;

Qy 1 MCGIVGIAQORDVAEILLEGRLREYRGYSAGLAVVDAEGHMTLRRLKGVQMLAAAE 60
Db 1 MCGIVGIAAERNITPILIEGLKREYRGYSAGVAVPDNEGRLQRCRVRGVKASLEGLA 60

Qy 61 EHPHGGTGAHTRWATHGEPSEVNAHPVSHIIVVHNGIIEHPELREELKARGYTFV 119
Db 61 GTPLRLGLAHTRWATHGAPTGNNAHPSSDEVAVVHNGIIEHPELRLKGLGYVP 120

Qy 120 VSETDTEVIAHLVNWELKOGGTLREAVLRAIPOLRGAYGTVMDSRHPDPTLLAARSGPL 179
Db 121 TSQTDTEVIAHLVNWELKOGGTLREAVLRAIPOLRGAYGTVMDSRHPDPTLLAARSGPL 180

Qy 180 VIGLGMENFTASDQALLPVTFRFLEEGDIAETRRSVNIYFDKTAEGVKRODIESNL 239
Db 181 VIGLGMENFTASDQALLPVTFRFLEEGDIAETRRSVNIYFDKTAEGVKRODIESNL 240

Qy 240 QYDAGDKGIYRHYMQEIEYQPNNAIKNTLTGRISHGQVDSLSELGPNADLLSKVEHIQIL 299
Db 241 GABAADKGRHYRHYMQEIEYQPNNAIKNTLTGRISHGQVDSLSELGPNADLLSKVEHIQIL 299

Qy 300 ACCTSNGMVSRYWFESLAGIPCDVEIASBFRYKSAVRNLSMITLSQSGETADTLAAGL 359
Db 301 ACCTSNGMVSRYWFESLAGIPCDVEIASBFRYKSAVRNLSMITLSQSGETADTLAAGL 360

Qy 360 LRUSKELGYLGLSLAICNVPGSSLVRESDLAMTNAGTEIGVASTKAFPTTQLTVLLMLVAK 419
Db 361 LRUSKELGYLGLSLAICNVPGSSLVRESDLAMTNAGTEIGVASTKAFPTTQLTVLLMLVAK 420

Qy 420 LSRL-LKGLDASIEHDIVHGLQALPSRIEOMLSODKRIEALAEFSDKHAFPLGRGQOY 478
Db 421 IGQVQRLADGVAEILVHGLQALPSRIEOMLSODKRIEALAEFSDKHAFPLGRGQOY 480

Qy 479 IALEGALKKEISYTHAEYAAAGELKHGPLALIDADMPIVIVVAPNNELLEKLKSNIEEVR 538
Db 481 VALEGALKKEISYTHAEYAAAGELKHGPLALIDADMPIVIVVAPNNELLEKLKSNIEEVR 540

Qy 539 ARGQQLYVAFADQAGFVSSDNMHIEMPHVEEVIAPITVPLQLLAYHVALIKGTVDQOP 598
Db 541 ARGQQLYVAFADQAGFVSSDNMHIEMPHVEEVIAPITVPLQLLAYHVALIKGTVDQOP 600

Qy 599 RNLAKSVTVE 609
Db 601 RNLAKSVTVE 611

RESULT 9
A84933
Glutamine-fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) [imported] - B
N;Alternate names: D-fructose-6-phosphate amidotransferase
C;Species: Buchnera sp.
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 28-Jul-2003
C;Accession: A84933
R;Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.
Nature 407, 81-86, 2000
A;Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. A
A;Reference number: A84930; PMID:10993077
A;Accession: A84933
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-621 <STO>
A;Cross-references: UNIPARC:UPI000005E424; GB:AP0000398; GSPDB:GN00144
A;Experimental source: strain APS
C;Genetics:
A;Gene: glms; BU026
C;Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing)
C;Keywords: aminotransferase; isomerase
Query Match 60.5%; Score 1868; DB 2; Length 621;

Best Local Similarity 59.1%; Pred. No. 8.8e-113;
Matches 361; Conservative 116; Mismatches 130; Indels 4; Gaps 3;
QY 1 MCGIVGATAQDVAAIILEGLRRLRYGYDSAGLAVDAEGHMTLRLRLGKVMQLAQAAE 60
DB 13 MCGIVAAVTOINIANFLIDGKLEKLYRGYDSGLAVIDNKNIVIRICVGNELIKKTN 72
QY 61 EHLPLGGGTGIAHTRWATHGEPSEVNAHPHSEHIVVHNGIENHPELREELKARGYTFV 120
DB 73 KKKILSGVAVHTRWATHGKSKENTPHISSNIIVVHNGIENNSLRLGFLKKQGYFS 132
QY 121 SETDTEVIAHLVNNEL-KQGGTLREAVLRAIPQLRGAYGTIVMDSRHPDITLLAARGSP 179
DB 133 SDTDEVIHLHWEQNKXDSLKVQNSIKLDDGNSVMVIDQNNPSKLIARAAGSP 192
QY 180 VIGLGMENFIASQDALLPVRPFIIFLEGDIAETTRSVNIPDKTGAEVKQDIESNL 239
DB 193 IIGLGTENFIASQDIALHVTKRFIYLEGDIAIIVARKEINIFNKNNSIIQREEVVNSI 252
QY 240 QYDAGDKGIVRYHWOKEIYEQPNAIKNTLTGRISH-CQVDLSLGNPADELKSKVEHIOI 298
DB 253 EYSAKKGKRYNKEIHOQPSIRNTLKNRLTNSKNHFSSELGSKENNIFNTYHEHIOI 312
QY 299 LACTSYNSGMVSYWFESLAGIPCDVEIASFEFRYKSAVRNLSLMTITLSQSGETADTILA 358
DB 313 VACGTSYNAAMVSYWFEELANTIPCDVEIASFEFSRKLVRKSLITLITLSQSGETADTILS 372
QY 359 GLRLSKELGYLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFPTQLTVLLMLVA 418
DB 373 ALRYSKKLGYLGNLTICNMKSSSLVRESDFYITKAGLEIGVASTKFTTQLTVLLMLVA 432
QY 419 KLSRLKGLDASIEHDIYHGLQALPSRIEOMLSQDKRIEALAEFSDKHALLFLGRGDQYP 478
DB 433 KIINSKKNNTSKRIVQTLISILPVRLEELKXQILQDWTANLAKKNMFLGRGNQRP 492
QY 479 IALEGALKKEISYTHAEVAAAGELKHGPLALIDADMPVIVVAPNNELKLSNIEEVR 538
DB 493 IAMEGALKKEISYTHAEVAPSGELKHGPLALIDKNIPVIMTAPENSLLEKKNKEIC 552
QY 539 ARGQQLVFPADQAGFVSSDNMHIIEMPHVEEVIAPFIVTVPLQLLAYHVALIKGTDVDO 598
DB 553 SRGIVTVFSNQ8--FDYENINTIKLPYVEELIAPFIVTIPQLQFAYYVALKGRDIDQ 610
QY 599 PRNLAKSVTVE 609
DB 611 PRHLAKSVTVE 621

RESULT 10

E71272
glutamine-fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) glms - syphilis
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
C:Accession: E71272
R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A:Reference number: A71250; MUID:98332770; PMID:9665876
A:Accession: E71272
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-635 <COL>
A:Cross-references: UNIPROT:083833; UNIPARC:UPI0000164AEC; GB:AE001256; GB:AE000520; NID
A:Experimental source: strain Nichols
A:Genetics:
A:Gene: TP0861
C:Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing)
C:Keywords: aminotransferase; isomerase
F:2-635/Product: glutamine-fructose-6-phosphate transaminase (isomerizing) #status predi
F:2/Active site: Cys #status predicted

Query Match 55.2%; Score 1702; DB 2; Length 635;
Best Local Similarity 54.0%; Pred. No. 4.7e-102;
Matches 343; Conservative 98; Mismatches 168; Indels 26; Gaps 5;
QY 1 MCGIVGATAQDVAAIILEGLRRLRYGYDSAGLAVDAEGHMTLRLRLGKVMQLAQAAE 60
DB 1 MCGIVGMVAGRVDSGLLEGLRRLRYGYDSAGIIVAGSDCALRLRCRGEVQSLCALIG 60
QY 61 EHLPLGGGTGIAHTRWATHGEPSEVNAHPHSEHIVVHNGIENHPELREELKARGYTFV 120
DB 61 QSPLCGTGIAHTRWATHGKPCAAAHPCSESVAIVHNGIVENHRSIREMLVTRGYFFH 120
QY 121 SETDTEVIAHLVNNELKOGGTLREAVLRAIPQLRGAYGTIVMDSRHPDITLLAARGSP 180
DB 121 SQTDEVLHLLHWEIYTHALLAVKVLTVQVRGTGYLLCWDAAASPGRLIAARGSP 180
QY 181 IGLGMENFIASQDALLPVRPFIIFLEGDIAETTRSVNIPDKTGAEVKQDIESNLQ 240
DB 181 VGLGCGENFVSDPLALAHVTRQRFIYLEGDIAVHRDVCVCHDAQGNVVARPVVYQMQ 240
QY 241 YDAGDKGIYRYHMQKEIYEQPNAIKNTLTGRISHGQVDLSL-----GPN-----D 287
DB 241 LCTQDKGTHRHMHQEIWQPHAIRHTLNAVMSFSSSSRAQVRTFGEDRVLDTGTSCKTFE 300
QY 288 ELLSKVEHIOILACGTSYNSGMVSYWFESLAGIPCDVEIASFEFRYKSAVRNLSLMTIL 347
DB 301 RLFRITRVRRIIACGTSYHAGLVARYWFEAPAGVGCQVEIASEYRYRTSVVHAREIVLTI 360
QY 348 QSGETADTLAHLRLSKELGYLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKRAFT 407
DB 361 QSGETADTLAHLRLAKTQGYLCALICNGARSLTVRESDAVLTHAGSELGVASTKFT 420
QY 408 TQLTVLLMLVAKLSRLKG-LDASIEHDIYHGLQALPSRIEOMLSQDKRIEALAEFSDKH 466
DB 421 TQLVCLLVLTIRIAQAQKILITQEPEDALSAAALQRLPQDVEHVLCEADVARCARHFVHAQ 480
QY 467 HALFLGRGDQVPALLEGALKKEISYTHAEVAAAGELKHGPLALIDADMPVIVVAPNN- 525
DB 481 HALFLGRGELYPITAEISALKKEISYTHAEVAAAGELKHGPLALVDQMPVVAIPASPG 540
QY 526 -LLEKLSNIEEVRARGQLYVFAD-----QDAGFVSSDNMHIIEMPHVEEVIAP 574
DB 541 VLFEKMASNIEEVRARGMLYIFTDVPERFGPCTPEADAPGACSQIVTVPSVSLTAP 600
QY 575 IFYTVPLQLLAYHVALIKGTDVQDQPRNLAKSVTVE 609
DB 601 IFYAVPLQLLAYHVALIKGTDIDQPRNLAKSVTVE 635

RESULT 11

A82844
glutamine-fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) [similarity] - Xy
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: A82844
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequenc
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: A82844
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-609 <SIM>
A:Cross-references: UNIPROT:Q9PH05; UNIPARC:UPI0000165AA4; GB:AE003867; GB:AE003849; NID:
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; AJ
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H.
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Praga, J.S.; Franca, S.C.; Franco, M.C.; Prohme
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuranae, E.E.; Laigre
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.

A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XP0141
C;Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing)
C;Keywords: aminotransferase; isomerase
F;2-609/Product: glutamine-fructose-6-phosphate transaminase (isomerizing) #status predi
F;2/Active site: Cys #status predicted

```
Query Match          54.9%; Score 1695; DB 2; Length 609;
Best Local Similarity 54.9%; Pred. No. 1.3e-101;
Matches 335; Conservative 103; Mismatches 170; Indels 2; Gaps 2;

Qy 1 MCGIVGAIARDVAEILLEGRLRYGYDSAGLAVVDAEGHMTLRRLGKVMQAQAAE 60
Db 1 MCGIVGAIARDVVPVLEGLKRLRYGYDSAGLAVVDAEGHMTLRRLGKVMQAQAAE 59

Qy 61 EPHLHGCTGTAHTRWATHGEPSEVNAHPHYSVHVVHNGIIEHNEPRLBELKARGYTFV 120
Db 60 QEGFTASLGIGHTRWATHGEGVTEANAHPHYSVHGVVLVHNGIIEHNEVQERLSALGVFQ 119

Qy 121 SETDTEVIAHLVNNELKQGGTLREAVLRATPOLRGAYGTVMDSRHPDTLLAARSGSPLV 180
Db 120 SQTDTDEVIAHLHYHMQGDDLLGALQCAVKALGTIYALAVMSAEFPCARMGCPLL 179

Qy 181 IGLGMGNFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIFDKTGAEVKRODIESNLQ 240
Db 180 IIGDGEHLVASDISAVIQATQVIFLEDGDTAEIRRDGISIFNAEQCPVERPLHLSNVS 239

Qy 241 YDAGDKGIYRHYMQKEIYEQNAIKNTLTGRISHGQVDLSLGNVADELKSKVHHQILA 300
Db 240 LSSLELGEFRHFMQKEIHEQPRVLADTMEAAIDAAGPPMLFGAQAESVFRGITGIQILA 299

Qy 301 CGTSYNGMSVRYWFESLAGIPCDVEIASFPRYKRSVRRNSLMTLSQSGETADTLAGL 360
Db 300 CGTSYAGLTARYWIEAIGLPCHVEIASERYKRYKAVYNQHLVVTISQSGETLDTLEAL 359

Qy 361 RLSKELGYLGSALCNVPGSSLVRESDLALMTNAGTEIGVASTKAFDTTQLTVLLMLVAKL 420
Db 360 KYAKALGHRHTLSICNAPDSAPRISELICVTRAGPEIGVASTKAFDTTQLVWLFLQALVAL 419

Qy 421 SRLKG-LDASIEHDIVGLQALPSRIEQLMSQDKRIEALAEFSDKHALFLRGDQYPI 479
Db 420 GVLRGAVDAEHAAYLEQLRQLPCGVQOALNLEPQIAAWAECSFASRHHALFLRGHLHYPI 479

Qy 480 ALEGALKKEISYTHAEAYAGELKHGKLPLALIDADMPVIVVAPNNELLEKLKSNIEVRA 539
Db 480 ALEGALKKEISYTHAEAYAGELKHGKLPLALVADMPVIVVAPNDSLEKVKSNMQSVRA 539

Qy 540 RGGQLYVFADQDAGFVSDNNHIIEMPHVEVIAPIFYTVPLQLLAYHVALIKGTVDQDP 599
Db 540 RGGELFVADQDSHFSEGLHVIRTLRHGTGVLSPVLVTIPVQLLAYHTALVRGTVDQKP 599

Qy 600 RNLAKSVTVE 609
Db 600 RNLAKSVTVE 609
```

RESULT 12

T45493
glutamine-fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) glms [similarity
N;Alternate names: glucosamine-6-phosphate synthase
C;Species: Thiobacillus ferrooxidans
C;Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 09-Jul-2004
C;Accession: T45493
R;Oppon, J.C.; Sarnovsky, R.J.; Craig, N.L.; Rawlings, D.E.
J. Bacteriol. 180, 3007-3012, 1998
A;Title: A Tn7-like transposon is present in the glms region of the obligately chemoaut

A;Reference number: Z22992; MUID:98269023; PMID:9603897
A;Accession: T45493
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-611 <OPP>
A;Cross-references: UNIPROT.Q56275; UNIPARC:UIP000016E22C; EMBL.AF032884; NID:g2653994;
A;Experimental source: ATCC 33020
C;Genetics:
A;Gene: glms
C;Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing)
C;Keywords: aminotransferase; isomerase

```
Query Match          51.7%; Score 1597; DB 2; Length 611;
Best Local Similarity 52.8%; Pred. No. 2.7e-95;
Matches 325; Conservative 101; Mismatches 179; Indels 10; Gaps 5;

Qy 1 MCGIVGAIARDVAEILLEGRLRYGYDSAGLAVVDAEGHMTLRRLGKVMQAQAAE 60
Db 1 MCGIVGVSKTDLVPMILEGLQRLRYGYDSAGLAILGADADLLRVRSVGRVAELTAADV 60

Qy 61 EPHLHGCTGTAHTRWATHGEPSEVNAHPHYS-BHI VVVHNGIIEHNEPRLBELKARGYTF 119
Db 61 ERGLQGVGIGHTRWATHGEGVRECNAHPMISHQIADVHNGIIEHNEPRLBELKARGYTF 120

Qy 120 VSETDTEVIAHLVNNELKQGGTLREAVLRATPOLRGAYGTVMDSRHPDTLLAARSGSPL 179
Db 121 TSETDTEVIAHLVHHVYRQTAPDLFAATRRRAVGLRGAYATAVISSGDPETVCVARMGCPL 180

Qy 180 VIGLGMGNFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIFDKTGAEVKRODIESNL 239
Db 181 LLGVADDGHYFASDVAALEFVTRRVLYLEDGDVAMLQRTLRITDQAGASRQREEHWSQL 240

Qy 240 QYDAGDKGIYRHYMQKEIYEQNAIKNTLTGRISHGQVDLSL-GPNADELKSKVHHQIOI 298
Db 241 SAAAVDLGPRHFMQKEIHEQPRAVADTLEGAL-NSQLDLDLWDGGAAMFRDVRDVLFP 299

Qy 299 LACGTSYNGMSVRYWFESLAGIPCDVEIASFPRYKRSVRRNSLMTLSQSGETADTTLA 358
Db 300 LASGTSYATVLVGRQWVESIVGIPAQAELGHEVRYRDSIPDPRLVVTLSQSGETLDTFE 359

Qy 359 GLRSKELGYLGSALCNVPGSSLVRESDLALMTNAGTEIGVASTKAFDTTQLTVLLMLVA 418
Db 360 ALBRKDLGLTRTLAICNVAESAIPRASALRFLTRAGPEIGVASTKAFDTTQLAALYLLAL 419

Qy 419 KLSRLKGLDASIEHDIVHG---LQALPSRIEQLMSQDKRIEALAEFSDKHALFLRG 474
Db 420 SLAKAPG---ASERCAGGSPGRLQLPGSVQHALLNLEPQLQGWAAFPASKDHFLRG 476

Qy 475 DQYPIALEGALKKEISYTHAEAYAGELKHGKLPLALIDADMPVIVVAPNNELLEKLKSN 534
Db 477 LHYPIALEGALKKEISYTHAEAYAGELKHGKLPLALVDRMPVVVVIAPNDRLLLEKLAAN 536

Qy 535 EYVARGGQLYVFADQDAGFVSDNNHIIEMPHVEVIAPIFYTVPLQLLAYHVALIKGT 594
Db 537 QEYHARGGELYVFADSDSHFNAGSVHVMRLPRHAGLLSPVHAIPVQLLAYHAALVKGT 596

Qy 595 DVDQPRNLAKSVTVE 609
Db 597 DVDPRNLAKSVTVE 611
```

RESULT 13

H82022
glutamine-fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) NMA0276 [similar
N;Alternate names: glucosamine fructose-6-phosphate aminotransferase [minomer]
C;Species: Neisseria meningitidis
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C;Accession: H82022
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jørgensen, K.; Leather, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis 22491.
A;Reference number: A81775; MUID:20222556; PMID:10761919

A:Accession: H82022
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-612 <PAR>
A:Cross-references: UNIPROT:Q9JWN9; UNIPARC:UPI0000165682; GB:AL162752; GB:AL157959; NID
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: glmS; NMA0276
C:Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing)
C:Keywords: aminotransferase; isomerase
F:2-612/Product: glutamine-fructose-6-phosphate transaminase (isomerizing) #status predicted
F:2/Active site: Cys #status predicted

Query Match 50.9%; Score 1570.5; DB 2; Length 612;
Best Local Similarity 52.8%; Pred. No. 1.4e-93;
Matches 325; Conservative 104; Mismatches 177; Indels 9; Gaps 9;

QY 1 MCGIVGAI-AQRDVAEILLEGRLLEYRGYDSAGLAVVDAGEHMTRLRLRGKVMQAQA 59
DB 1 MCGIVGAI-RAHHNVVDFLTGKRLLEYRGYDSSGIA-VNTDGKIKRVRRVGRVQLMEDAA 59

QY 60 EEHPLHGCTGTAHTRWATHGSPSEVNAHPHVS-EHIVVHNGIIEHPEELKARGYT 118
DB 60 REKISGGIGIGHRTWATHGVTPEPNAPHISGGMIAVHNGIIEHPESEKRLGLELYR 119

QY 119 FVSETDTEVIAHLVNMELKQ-GGTLREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGS 177
DB 120 FESQTDTEVIAHSINHEVAQNGKLFPAVQEAQVRFHGAIAIAVIAQDKPELVVARMGC 179

QY 178 PLVTGLMGENFIASDQALLPVTRRFIFLEGDIABITRRSV-NIPDKTGAEVKRDIE 236
DB 180 PLLVALGDDETFIASDVSAVIAFTRRVAYLEDGDIALASDGIKRLTDKSLPAERKVKV 239

QY 237 SNLQYDAGDKGIYRHYMKEIYEQPNALKNLTLCRISHGQVDLSLGNADLLSKVEHI 296
DB 240 SELSLALELGPYSHFQKEIHEQPRAIADTAEVFLDGGFIP-ENFGKNAKSVFESIRSV 298

QY 297 QILACGTSYNGMVSRYWFESLAGIPCDVEIASFEFRYKSAVRNLSMITLSQSGTADT 356
DB 299 KILACGTSYVAALTAKYWLSEIAKIPSDVEIASERYSVIADPDQVLVITISQSGETLDT 358

QY 357 LAGRLSKEILGYLSLAICNVPSSLVRESDLALMTNAGTEIGVASTKFTTQLTVLLML 416
DB 359 MEALKYAKSLGHRHSLISICNVMESALPRESSLVLYTRAGAEIGVASTKFTTQLVALFGL 418

QY 417 VAKLSRLKGL-DASIEHDI VHGLQALPSRIEQMLSDKRIEALAEFSDKHHLFLRGD 475
DB 419 AVTLAKVRLGVSEDEARYTEELRQLPGSVQHALLNDEPQIAAWAQFAKTSALFLRGI 478

QY 476 QYPALLEGALKKEISYTHAEYAAGELKHGPIALIDAMPVIVVAPNNELLEKLKSNI 535
DB 479 HYPALLEGALKKEITYTHAEYAPAGELKHGPLALVDENMPVVIAPNDSLLDKVKANMQ 538

QY 536 EVPARGOLVVFADQAGFVSSDNMHIEMP-HVEVIAPIFYTPVPLQLLAHVHALIKGT 594
DB 539 EVGARGELVFADLDSNFNATEGVHVIAPRPHVGE-LSPVVHTIPVQLLSYHVALRG 597

QY 595 DVDQPRNLAKSVTV 609
DB 598 DVDKPRNLAKSVTIVE 612

RESULT 14
B81246
N:Alternate names: glucosamine fructose-6-phosphate aminotransferase [isomerase]
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: B81246
R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
xi, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masiugnani, V.; Pizza, M.
Science 287, 1809-1915, 2000

A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ver
A:Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.
A:Reference number: AB1000; MUID:20175755; PMID:10710307
A:Accession: B81246
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-612 <TET>
A:Cross-references: UNIPROT:Q9KIP9; UNIPARC:UPI000016563D; GB:AE002361; GB:AE002098; NID
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB0031
C:Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing)
C:Keywords: aminotransferase; isomerase
F:2-612/Product: glutamine-fructose-6-phosphate transaminase (isomerizing) #status predicted
F:2/Active site: Cys #status predicted

Query Match 50.8%; Score 1566.5; DB 2; Length 612;
Best Local Similarity 52.3%; Pred. No. 2.5e-93;
Matches 321; Conservative 104; Mismatches 182; Indels 7; Gaps 7;

QY 1 MCGIVGAI-AQRDVAEILLEGRLLEYRGYDSAGLAVVDAGEHMTRLRLRGKVMQAQA 59
DB 1 MCGIVGAI-RAHHNVVDFLTGKRLLEYRGYDSSGIA-VNTDGKIKRVRRVGRVQLMEDAA 59

QY 60 EEHPLHGCTGTAHTRWATHGSPSEVNAHPHVS-EHIVVHNGIIEHPEELKARGYT 118
DB 60 REKISGGIGIGHRTWATHGVTPEPNAPHISGGMIAVHNGIIEHPESEKRLGLELYR 119

QY 119 FVSETDTEVIAHLVNMELKQ-GGTLREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGS 177
DB 120 FESQTDTEVIAHSINHEVAQNGKLFPAVQEAQVRFHGAIAIAVIAQDKPELVVARMGC 179

QY 178 PLVTGLMGENFIASDQALLPVTRRFIFLEGDIABITRRSV-NIPDKTGAEVKRDIE 236
DB 180 PLLVALGDDETFIASDVSAVIAFTRRVAYLEDGDIALASDGIKRLTDKSLPAERKVKV 239

QY 237 SNLQYDAGDKGIYRHYMKEIYEQPNALKNLTLCRISHGQVDLSLGNADLLSKVEHI 296
DB 240 SELSLALELGPYSHFQKEIHEQPRAIADTAEVFLDGGFIP-ENFGKNAKSVFESIRSV 298

QY 297 QILACGTSYNGMVSRYWFESLAGIPCDVEIASFEFRYKSAVRNLSMITLSQSGTADT 356
DB 299 KILACGTSYVAALTAKYWLSEIAKIPSDVEIASERYSVIADPDQVLVITISQSGETLDT 358

QY 357 LAGRLSKEILGYLSLAICNVPSSLVRESDLALMTNAGTEIGVASTKFTTQLTVLLML 416
DB 359 MEALKYAKSLGHRHSLISICNVMESALPRESSLVLYTRAGAEIGVASTKFTTQLVALFGL 418

QY 417 VAKLSRLKGL-DASIEHDI VHGLQALPSRIEQMLSDKRIEALAEFSDKHHLFLRGD 475
DB 419 AVTLAKVRLGVSEDEARYTEELRQLPGSVQHALLNDEPQIAAWAQFAKTSALFLRGI 478

QY 476 QYPALLEGALKKEISYTHAEYAAGELKHGPIALIDAMPVIVVAPNNELLEKLKSNI 535
DB 479 HYPALLEGALKKEITYTHAEYAPAGELKHGPLALVDENMPVVIAPNDSLLDKVKANMQ 538

QY 536 EVPARGOLVVFADQAGFVSSDNMHIEMP-HVEVIAPIFYTPVPLQLLAHVHALIKGT 595
DB 539 EVGARGELVFADLDSNFNATEGVHVIAPRPHVGLSPVVHTIPVQLLAHYHTALARGTD 598

QY 596 DVDQPRNLAKSVTV 609
DB 599 DVDKPRNLAKSVTIVE 612

RESULT 15
AD3595
glutamine-fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) [imported] - Bru
C:Species: *Brucella melitensis*
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C:Accession: AD3595
R:DeiVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, I
; Mazur, M.; Goitsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letesc

Search completed: August 7, 2006, 09:25:39
Job time : 27 secs

This Page Blank (uspto)

RC STRAIN=K-12;
RX PubMed=16397293; DOI=10.1093/nar/gkj150;
RA Riley M., Abe T., Arnaud M.B., Berlyn M.K., Blattner F.R.,
RA Chaudhuri R.R., Glaser J.D., Horiuchi T., Keseler I.M., Kosuge T.,
RA Mori H., Perna N.T., Plunkett G. III, Rudd K.E., Serres M.H.,
RA Thomas G.H., Thomson N.R., Wishart D., Wanner B.L.;
RT "Escherichia coli K-12: a cooperatively developed annotation snapshot-
RT -2005";
RL Nucleic Acids Res. 34:1-9(2006).
RN [7]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=97349980; PubMed=9205837; DOI=10.1093/dnares/4.2.91;
RA Yamamoto Y., Aiba H., Baba T., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nishimoto H.,
RA Oshima T., Oyama S., Saito N., Sampei G., Satoh Y., Sivasundaram S.,
RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
RA Yamagata S., Horiuchi T.;
RT "Construction of a contiguous 874-kb sequence of the Escherichia coli
RT -K12 genome corresponding to 50.0-68.8 min on the linkage map and
RT analysis of its sequence features";
RL DNA Res. 4:91-113(1997).
RN [8]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=97251358; PubMed=9097040; DOI=10.1093/dnares/3.6.379;
RA Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA Kasai H., Kimura S., Kitakawa M., Kitagawa M., Makino K., Miki T.,
RA Mizobuchi K., Mori H., Mori T., Motomura K., Nakade S., Nakamura Y.,
RA Nishimoto H., Nishio Y., Oshima T., Saito N., Sampei G., Seki Y.,
RA Sivasundaram S., Tagami H., Takeda J., Takemoto K., Wada C.,
RA Yamamoto Y., Horiuchi T.;
RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 40.1-50.0 min region on the linkage map";
RL DNA Res. 3:379-392(1996).
RN [9]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=97251357; PubMed=9097039; DOI=10.1093/dnares/3.6.363;
RA Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T., Saito N.,
RA Sampei G., Seki Y., Sivasundaram S., Tagami H., Takeda J.,
RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 28.0-40.1 min region on the linkage map";
RL DNA Res. 3:363-377(1996).
RN [10]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=97094878; PubMed=8940112; DOI=10.1074/jbc.271.49.31145;
RA Arn E.A., Abelson J.N.;
RT "The 2'-5' RNA ligase of Escherichia coli. Purification, cloning, and
RT genomic disruption";
RL J. Biol. Chem. 271:31145-31153(1996).
RN [11]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=97061202; PubMed=8905232; DOI=10.1093/dnares/3.3.137;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Mizobuchi K.,
RA Mori H., Moromura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map";
RL DNA Res. 3:137-155(1996).
RN [12]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=94261430; PubMed=8202364;
RA Fujita N., Mori H., Yura T., Ishihama A.;
RT "Systematic sequencing of the Escherichia coli genome: analysis of the
RT 2.4-4.1 min (110,917-193,643 bp) region";
RL Nucleic Acids Res. 22:1637-1639(1994).
RN [13]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=94240115; PubMed=8183897;
RA Janosi L., Shimizu I., Kaji A.;
RT "Ribosome recycling factor (ribosome releasing factor) is essential
RT for bacterial growth";
RL Proc. Natl. Acad. Sci. U.S.A. 91:4249-4253(1994).
RN [14]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=94124004; PubMed=7904973; DOI=10.1016/0378-1119(93)90470-N;
RA Allikmets R., Gerrard B.C., Court D., Dean M.C.;
RT "Cloning and organization of the abc and mdl genes of Escherichia
RT coli: relationship to eukaryotic multidrug resistance";
RL Gene 136:231-236(1993).
RN [15]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=94018640; PubMed=8412694;
RA van Heeswijk W.C., Rabenberg M., Westerhoff H.V., Kahn D.D.;
RT "The genes of the glutamine synthetase adenylation cascade are not
RT regulated by nitrogen in Escherichia coli";
RL Mol. Microbiol. 9:443-458(1993).
RN [16]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=93259920; PubMed=8387890;
RA Zhao S., Sandt C.H., Feulner G., Vlazny D.A., Gray J.A., Hill C.W.;
RT "Rhs elements of Escherichia coli K-12: complex composites of shared
RT and unique components that have different evolutionary histories";
RL J. Bacteriol. 175:2799-2808(1993).
RN [17]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=93123180; PubMed=8419307;
RA Yamada M., Asaoka S., Saier M.H. Jr., Yamada Y.;
RT "Characterization of the gcd gene from Escherichia coli K-12 W3110 and
RT regulation of its expression";
RL J. Bacteriol. 175:568-571(1993).
RN [18]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=93116053; PubMed=1474579;
RA Cormack R.S., Mackie G.A.;
RT "Structural requirements for the processing of Escherichia coli 5 S
RT ribosomal RNA by RNase E in vitro";
RL J. Mol. Biol. 228:1078-1090(1992).
RN [19]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=93094132; PubMed=1459951;
RA Gervais F.G., Drapeau G.R.;
RT "Identification, cloning, and characterization of rcsF, a new
RT regulator gene for exopolysaccharide synthesis that suppresses the
RT division mutation fts284 in Escherichia coli K-12";
RL J. Bacteriol. 174:8016-8022(1992).
RN [20]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;
RX MEDLINE=93077430; PubMed=1447125;
RA Yamanaka K., Ogura T., Niki H., Hiraga S.;
RT "Identification and characterization of the smbA gene, a suppressor of
RT the mukB null mutant of Escherichia coli";
RL J. Bacteriol. 174:7517-7526(1992).
RN [21]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K-12;

| | |
|--|--|
| RX | MEDLINE=93011013; PubMed=1396599; |
| RA | Condon C., Phillips J., Fu Z.Y., Squires C., Squires C.L.; |
| Query Match 100.0%; Score 3086; DB 2; Length 609; | |
| Best Local Similarity 100.0%; Pred. No 2.8e-186; | |
| Matches 609; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | |
| Qy | 1 MCGIVGAIAQDVAEILLEGRLRLEYRGYDSAGLAVDAAGHMTLRRLRGVQMLAAAE 60 |
| Db | 1 MCGIVGAIAQDVAEILLEGRLRLEYRGYDSAGLAVDAAGHMTLRRLRGVQMLAAAE 60 |
| Qy | 61 EPHLGCGTGAHTRWATHGEPSEVNAHPHVSEHIVVHNGIIEHNEPRLBELKARGYTFV 120 |
| Db | 61 EPHLGCGTGAHTRWATHGEPSEVNAHPHVSEHIVVHNGIIEHNEPRLBELKARGYTFV 120 |
| Qy | 121 SETDTEVIAHLVNWELKOGGTLRAVLRAIPQLRGAYGTVMDSRHPDTLAARSGSPLV 180 |
| Db | 121 SETDTEVIAHLVNWELKOGGTLRAVLRAIPQLRGAYGTVMDSRHPDTLAARSGSPLV 180 |
| Qy | 181 IGLGMGFNFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIFDKTGAEVKRDQIESNLQ 240 |
| Db | 181 IGLGMGFNFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIFDKTGAEVKRDQIESNLQ 240 |
| Qy | 241 YDAGDKGIYRHYMKETIEYQPNAINKNTLTGRISHGQVDLSLGNADDELLSKVEHIQILA 300 |
| Db | 241 YDAGDKGIYRHYMKETIEYQPNAINKNTLTGRISHGQVDLSLGNADDELLSKVEHIQILA 300 |
| Qy | 301 CGTSYNGSMVSRVWFESLAGI PCDVEIASEFRYKRSVAVRNSLMTILTSQSGETADTLA 360 |
| Db | 301 CGTSYNGSMVSRVWFESLAGI PCDVEIASEFRYKRSVAVRNSLMTILTSQSGETADTLA 360 |
| Qy | 361 RLSKELGYLGLSLAICNVPGSSLVRESDLALMTNAGTIGVASTKFTTQLTVLLMLVAKL 420 |
| Db | 361 RLSKELGYLGLSLAICNVPGSSLVRESDLALMTNAGTIGVASTKFTTQLTVLLMLVAKL 420 |
| Qy | 421 SRLKGLDASTIEHDIVHGLQALPSRIEOMLSQDKRIEALAEFSDKXHALFLGRGDQYPIA 480 |
| Db | 421 SRLKGLDASTIEHDIVHGLQALPSRIEOMLSQDKRIEALAEFSDKXHALFLGRGDQYPIA 480 |
| Qy | 481 LEGALKLKETSITHAEAYAGELKHGPLALIDAMPVIVVAPNNLEKLSNIEVRAR 540 |
| Db | 481 LEGALKLKETSITHAEAYAGELKHGPLALIDAMPVIVVAPNNLEKLSNIEVRAR 540 |
| Qy | 541 GGOQYVFADQDAGFVSSDNNHIIEMPHVEVIAPIFTVPLQLLAYHVALIKGTVDVQPR 600 |
| Db | 541 GGOQYVFADQDAGFVSSDNNHIIEMPHVEVIAPIFTVPLQLLAYHVALIKGTVDVQPR 600 |
| Qy | 601 NLAQSVTVE 609 |
| Db | 601 NLAQSVTVE 609 |

RESULT 2

| | | | |
|------------|--|---------------------------------------|---------|
| GLMS_ECOLI | STANDARD; | PRT; | 608 AA. |
| ID | GLMS_ECOLI | | |
| AC | P17169; P76745; | | |
| DT | 01-AUG-1990, | integrated into UniProtKB/Swiss-Prot. | |
| DT | 15-DEC-1998, | sequence version 3. | |
| DT | 07-MAR-2006, | entry version 70. | |
| DE | Glucosamine--fructose-6-phosphate aminotransferase [isomerizing] | | |
| DE | (EC 2.6.1.16) (Hexosephosphate aminotransferase) (D-fructose-6- | | |
| DE | phosphate amidotransferase) (G6PT) (L-glutamine-D-fructose-6-phosphate | | |
| DE | amidotransferase) (Glucosamine-6-phosphate synthase). | | |
| GN | Name=glms; OrderedLocusNames=bb3729; | | |
| OS | Escherichia coli. | | |
| OC | Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; | | |
| OC | Enterobacteriaceae; Escherichia. | | |
| OX | NCBI_TaxID=562; | | |
| RN | [1] | | |
| RN | NUCLEOTIDE SEQUENCE [GENOMIC DNA]. | | |
| RX | MEDLINE=85121806; PubMed=6395859; | | |
| RA | Walker J.E., Gay N.J., Saraste M., Eberle A.N.; | | |
| RT | "DNA sequence around the Escherichia coli unc operon. Completion of | | |

| | |
|----|--|
| RT | the sequence of a 17 kilobase segment containing asnA, oriC, unc, glms |
| RT | and phos."; |
| RL | Biochem. J. 224:799-815(1984). |
| RN | [2] |
| RP | NUCLEOTIDE SEQUENCE [GENOMIC DNA]. |
| RC | STRAIN=K12 / MGI655; |
| RX | MEDLINE=93315143; PubMed=7686882; |
| RA | Burland V.D., Plunkett G. III, Daniels D.L., Blattner F.R.; |
| RT | "DNA sequence and analysis of 136 kilobases of the Escherichia coli |
| RT | genome: organizational symmetry around the origin of replication."; |
| RL | Genomics 16:551-561(1993). |
| RN | [3] |
| RP | PROTEIN SEQUENCE OF 48-51; 218-230; 488-492; 504-507 AND 600-608. |
| RX | MEDLINE=92007872; PubMed=1915361; |
| RA | Golinelli-Pimpaneau B., Badet B.; |
| RT | "Possible involvement of Lys603 from Escherichia coli glucosamine-6- |
| RT | phosphate synthase in the binding of its substrate fructose 6- |
| RT | phosphate."; |
| RL | Eur. J. Biochem. 201:175-182(1991). |
| RN | [4] |
| RP | NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 489-608. |
| RX | MEDLINE=82220022; PubMed=6283361; |
| RA | Lichtenstein C., Brenner S.; |
| RT | "Unique insertion site of Tn7 in the E. coli chromosome."; |
| RL | Nature 297:601-603(1982). |
| RN | [5] |
| RP | NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 606-608. |
| RX | MEDLINE=86215091; PubMed=3010949; |
| RA | Gay N.J., Tybulewicz V.L.-J., Walker J.E.; |
| RT | "Insertion of transposon Tn7 into the Escherichia coli glms |
| RT | transcriptional terminator."; |
| RL | Biochem. J. 234:111-117(1986). |
| RN | [6] |
| RP | NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 596-608. |
| RX | MEDLINE=88086894; PubMed=2826397; |
| RA | McKown R.L., Orle K.A., Chen T., Craig N.L.; |
| RT | "Sequence requirements of Escherichia coli attTn7, a specific site of |
| RT | transposon Tn7 insertion."; |
| RL | J. Bacteriol. 170:352-358(1988). |
| RN | [7] |
| RP | CHARACTERIZATION. |
| RX | MEDLINE=88281539; PubMed=3134953; DOI=10.1016/0300-9084(88)90073-9; |
| RA | Dutka-Malen S., Mazodier P., Badet B.; |
| RT | "Molecular cloning and overexpression of the glucosamine synthetase |
| RT | gene from Escherichia coli."; |
| RL | Biochimie 70:287-290(1988). |
| RN | [8] |
| RP | X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 1-240. |
| RX | MEDLINE=96434326; PubMed=8805567; DOI=10.1016/S0969-2126(96)00087-1; |
| RA | Iakupov M.N., Obmolova G., Butterworth S., Badet-Denisot M.-A., |
| RA | Badet B., Polikarpov I., Littlechild J.A., Teplyakov A.; |
| RT | "Substrate binding is required for assembly of the active conformation |
| RT | of the catalytic site in Ntn amidotransferases: evidence from the 1.8- |
| RT | A crystal structure of the glutaminase domain of glucosamine 6- |
| RT | phosphate synthase."; |
| RL | Structure 4:801-810(1996). |
| RN | [9] |
| RP | X-RAY CRYSTALLOGRAPHY (1.57 ANGSTROMS) OF 243-608. |
| RX | MEDLINE=98416699; PubMed=9739095; DOI=10.1016/S0969-2126(98)00105-1; |
| RA | Teplyakov A., Obmolova G., Badet-Denisot M.-A., Badet B., |
| RA | Polikarpov I.; |
| RT | "Involvement of the C terminus in intramolecular nitrogen channeling |
| RT | in glucosamine 6-phosphate synthase: evidence from a 1.6-A crystal |
| RT | structure of the isomerase domain."; |
| RL | Structure 6:1047-1055(1998). |
| RN | [10] |
| RP | X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 241-608. |
| RX | MEDLINE=99190083; PubMed=10091662; |
| RA | Teplyakov A., Obmolova G., Badet-Denisot M.A., Badet B.; |
| RT | "The mechanism of sugar phosphate isomerization by glucosamine 6- |
| RT | phosphate synthase."; |
| RL | Protein Sci. 8:596-602(1999). |
| CC | -!- FUNCTION: Catalyzes the first step in hexosamine metabolism, |

converting fructose-6P into glucosamine-6P using glutamine as a nitrogen source.
-|- CATALYTIC ACTIVITY: L-glutamine + D-fructose 6-phosphate = L-glutamate + D-glucosamine 6-phosphate.
-|- SUBUNIT: Homodimer.
-|- INTERACTION: P00957:alaS; N5Exp=1; IntAct=EBI-551022, EBI-544061; P08622:dnajJ; N5Exp=1; IntAct=EBI-551022, EBI-545285; P76552:euthJ; N5Exp=1; IntAct=EBI-551022, EBI-551031; P62615:apeE; N5Exp=1; IntAct=EBI-551022, EBI-562202; P61175:rpIv; N5Exp=1; IntAct=EBI-551022, EBI-542255; P21166:crkH; N5Exp=1; IntAct=EBI-551022, EBI-550268; P76093:yhbD; N5Exp=1; IntAct=EBI-551022, EBI-551038; P33366:yohD; N5Exp=1; IntAct=EBI-551022, EBI-551046; -|- SUBCELLULAR LOCATION: Cytoplasm.
-|- SIMILARITY: In the C-terminal section; belongs to the SIS family. GPAT subfamily.
-|- SIMILARITY: Contains 1 type-2 glutamine amidotransferase domain.

Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
Distributed under the Creative Commons Attribution-NoDerivs License

EMBL; X01631; CAA25785.1; -; Genomic_DNA.
EMBL; L10328; AAA62080.1; -; Genomic_DNA.
EMBL; U00096; AAC76752.1; -; Genomic_DNA.
EMBL; V00620; CAA23894.1; -; Genomic_DNA.
EMBL; M18980; AAA23836.1; -; Genomic_DNA.
PIR; B65176; XNECGM.
PDB; 1JXA; X-ray; A/B/C=1-608.
PDB; 1MOQ; X-ray; @=241-608.
PDB; 1MOR; X-ray; @=241-608.
PDB; 1MOS; X-ray; A=241-608.
PDB; 1XFF; X-ray; A/B=1-240.
PDB; 1XFG; X-ray; A/B=1-240.
PDB; 2BPJ; X-ray; A/B=1-608.
PDB; 2BPL; X-ray; A/B/C=1-608.
DR MEROPS; C44.971; -.
DR GenomeReviews; U00096_GR; b3729.
DR EcoGene; EB0377; -.
DR EcoGene; EG10382; glms.
DR BioCyc; EcoCyc:i-GLN-FRUCT-6-P-AMINOTRANS-MONOMER; -.
DR GO; GO:0005515; F:protein binding; IPI.
DR HAMAP; MF_00164; -; 1.
DR InterPro; IPR000583; Gataase_2.
DR InterPro; IPR005855; Glms_trans.
DR Pfam; PF00310; Gataase_2; 1.
DR Pfam; PF01380; SIS; 2.
DR TIGRFAMs; TIGR01135; glms; 1.
DR PROSITE; PS00443; GATASE_TYPE_II; 1.
KW 3D-structure; Aminotransferase; Complete proteome;
KW Direct protein sequencing; Glutamine amidotransferase; Transferase.
FT INIT_MET 0 0
FT CHAIN 1 608 Glucosamine--fructose-6-phosphate
FT aminotransferase [isomerizing].
FT /FTId=PRO_0000135328.
FT GATase.
FT Glutamine amidotransferase.
FT REGION 1 181
FT ACT_SITE 1 1
FT ACT_SITE 603 603
FT CONFLICT 418 419 Isomerization Fru-6P.
FT STRAND 2 7 KL -> NV (in Ref. 1).
FT STRAND 9 10
FT HELIX 13 26
FT STRAND 29 36
FT STRAND 38 39
FT STRAND 42 49
FT STRAND 51 60
FT STRAND 61 61
FT STRAND 66 73
FT STRAND 76 78
FT STRAND 80 81
FT TURN 82 84
FT STRAND 85 86

| | | | |
|----|--------|-----|-----|
| FT | STRAND | 88 | 90 |
| FT | TURN | 91 | 92 |
| FT | STRAND | 93 | 101 |
| FT | TURN | 102 | 103 |
| FT | HELI | 104 | 113 |
| FT | TURN | 114 | 115 |
| FT | STRAND | 120 | 120 |
| FT | STRAND | 123 | 123 |
| FT | HELI | 124 | 136 |
| FT | TURN | 137 | 137 |
| FT | STRAND | 138 | 139 |
| FT | HELI | 141 | 151 |
| FT | STRAND | 154 | 162 |
| FT | TURN | 163 | 164 |
| FT | TURN | 166 | 167 |
| FT | STRAND | 169 | 176 |
| FT | STRAND | 179 | 182 |
| FT | STRAND | 184 | 185 |
| FT | STRAND | 187 | 192 |
| FT | HELI | 193 | 195 |
| FT | TURN | 196 | 199 |
| FT | STRAND | 201 | 205 |
| FT | TURN | 208 | 209 |
| FT | STRAND | 211 | 214 |
| FT | STRAND | 216 | 217 |
| FT | STRAND | 219 | 222 |
| FT | TURN | 224 | 225 |
| FT | STRAND | 228 | 229 |
| FT | STRAND | 233 | 235 |
| FT | TURN | 246 | 247 |
| FT | STRAND | 249 | 250 |
| FT | HELI | 251 | 257 |
| FT | TURN | 258 | 258 |
| FT | HELI | 259 | 267 |
| FT | TURN | 268 | 269 |

Query Match 99.8%; Score 3081; DB 1; Length 608;
Best Local Similarity 100.0%; Pred. No. 5.8e-186;
Matches 508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|-----|--|-----|
| QY | 2 | CGIVGAIQRDVAEILLEGRLRLEYRGYDSAGLVDAEGHMTLRLRLGKVQMLAAAE | 61 |
| Db | 1 | CGIVGAIQRDVAEILLEGRLRLEYRGYDSAGLVDAEGHMTLRLRLGKVQMLAAAE | 60 |
| QY | 62 | HPLHGGTGIATRWATHGEPSEVNAHPVSEHIVVNHGIIENHEPLREELKARGYTFVS | 121 |
| Db | 61 | HPLHGGTGIATRWATHGEPSEVNAHPVSEHIVVNHGIIENHEPLREELKARGYTFVS | 120 |
| QY | 122 | ETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVINDSRHPDTLLAARSGSLVI | 181 |
| Db | 121 | ETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVINDSRHPDTLLAARSGSLVI | 180 |
| QY | 182 | GLGNGENFIASDQALALPVTTRRFIFLEEGDIAETIRRSVNI FDKTGAEVKQDIESNQY | 241 |
| Db | 181 | GLGNGENFIASDQALALPVTTRRFIFLEEGDIAETIRRSVNI FDKTGAEVKQDIESNQY | 240 |
| QY | 242 | DAGDKGIYRHYMQEIVEQPNAIKNTLTGRISHQVDLSELGPNADLLSVEHIQILAC | 301 |
| Db | 241 | DAGDKGIYRHYMQEIVEQPNAIKNTLTGRISHQVDLSELGPNADLLSVEHIQILAC | 300 |
| QY | 302 | GTSYNSGMVSRYPWFESLAGIPCDVEIASFEFRYKRSVAVRRNSLMTLSQSGETADTLA | 361 |
| Db | 301 | GTSYNSGMVSRYPWFESLAGIPCDVEIASFEFRYKRSVAVRRNSLMTLSQSGETADTLA | 360 |
| QY | 362 | LSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFPTQTLTVLLMLVAK | 421 |
| Db | 361 | LSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFPTQTLTVLLMLVAK | 420 |
| QY | 422 | RLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRLTEALAEFSDKHHLFLRGQOYPTAL | 481 |
| Db | 421 | RLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRLTEALAEFSDKHHLFLRGQOYPTAL | 480 |
| QY | 482 | EGALKLKEISYIHAEYAAAGELKHGLPLALIDADMPVIVVAPNNLELKLKSNIEVRARG | 541 |

```
|||||
Db 481 EGALKLKEISYTHAEAYAGELKHGFLALIDAMPVIVVAPNNLEKLNIEEVRARG 540
Qy 542 GOLVVFADQAGFVSSDNHIIEMPHVEEVIAPFYTVPLQLLAYHVALIKGTDVDDPRN 601
Db 541 GOLVVFADQAGFVSSDNHIIEMPHVEEVIAPFYTVPLQLLAYHVALIKGTDVDDPRN 600
Qy 602 LAKSVTVE 609
Db 601 LAKSVTVE 608

RESULT 3
Q329R8_SHIDS
ID Q329R8_SHIDS PRELIMINARY; PRT; 609 AA.
AC Q329R8;
DT 06-DEC-2005, integrated into UniProtKB/TrEMBL.
DT 06-DEC-2005, sequence version 1.
DE L-glutamine:D-fructose-6-phosphate aminotransferase.
GN Name=glms; OrderedLocusNames=SDY_4019;
OS Shigella dysenteriae serotype 1 (strain Sdi97).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=300267;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=16275786; DOI=10.1093/nar/gki954;
RA Yang F., Yang J., Zhang X., Chen L., Jiang Y., Yan Y., Tang X.,
RA Wang J., Xiong Z., Dong J., Xue Y., Zhu Y., Xu X., Sun L., Chen S.,
RA Nie H., Peng J., Xu J., Wang Y., Yuan Z., Wen Y., Yao Z., Shen Y.,
RA Qiang B., Hou Y., Yu J., Jin Q.;
RT "Genome dynamics and diversity of Shigella species, the etiologic
RT agents of bacillary dysentery."
RL Nucleic Acids Res. 33:6445-6458(2005).
CC -----
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL: CP000034; AB663937.1; -; Genomic_DNA.
DR GO: GO:0005737; C:cytoplasm; IEA.
DR GO: GO:0004360; F:glutamine-fructose-6-phosphate transaminase. . ; IEA.
DR GO: GO:0005529; F:sugar binding; IEA.
DR GO: GO:0016740; F:transferase activity; IEA.
DR GO: GO:0016051; P:carbohydrate biosynthesis; IEA.
DR GO: GO:0005975; P:carbohydrate metabolism; IEA.
DR GO: GO:0008152; P:metabolism; IEA.
KW Aminotransferase; Complete proteome; Transferase.
SQ SEQUENCE 609 AA; 66904 MW; 6373301513227484 CRC64;

Query Match 99.8%; Score 3081; DB 2; Length 609;
Best Local Similarity 99.8%; Pred. No. 5.8e-186;
Matches 608; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MCGIVGAIAORDVAEILLEGRLREYRGYSAGLAVVDAGHMTLRRLRGKVMQAAAE 60
Db 1 MCGIVGAIAORDVAEILLEGRLREYRGYSAGLAVVDAGHMTLRRLRGKVMQAAAE 60

Qy 61 EHPLHGCTGIAHTRWATHGSPSEVNAHPHVSEHIVVHNGIIEHNEHPLREBELKARGYTFV 120
Db 61 EHPLHGCTGIAHTRWATHGSPSEVNAHPHVSEHIVVHNGIIEHNEHPLREBELKARGYTFV 120

Qy 121 SETDTEVIAHVNWELKOGGTLREAVLRAPOLRGAYCTVIMDSRHPDPTLLAARSGSPLV 180
Db 121 SETDTEVIAHVNWELKOGGTLREAVLRAPOLRGAYCTVIMDSRHPDPTLLAARSGSPLV 180

Qy 181 IGLGMGFNTASDQALLPVTTRRIFLEEGDIAEITRRSVNIFDKTGAEVKRODIESNLQ 240
Db 181 IGLGMGFNTASDQALLPVTTRRIFLEEGDIAEITRRSVNIFDKTGAEVKRODIESNLQ 240

Qy 241 YDAGDKGIYRHYMQEYEQPNAIKNTLTGRISHGQVDLSLGNADLLSKVEHIQILA 300
Db 241 YDAGDKGIYRHYMQEYEQPNAIKNTLTGRISHGQVDLSLGNADLLSKVEHIQILA 300
```

```
Qy 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASFEFRYKRSAVRNSLMTLSQSGETADTLAGL 360
Db 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASFEFRYKRSAVRNSLMTLSQSGETADTLAGL 360
Qy 361 RLSKEILGYLGLSLAICNVPGSSLIARVRESDLALMTNAGTEIGVASTKAFPTOLTLLMLVAKL 420
Db 361 RLSKEILGYLGLSLAICNVPGSSLIARVRESDLALMTNAGTEIGVASTKAFPTOLTLLMLVAKL 420
Qy 421 SRLKGLDASIEHDI VHGLQALPSRIEOMLSQDKRIEALAEFSDKHHALFLRGDQYPIA 480
Db 421 SRLKGLDASIEHDI VHGLQALPSRIEOMLSQDKRIEALAEFSDKHHALFLRGDQYPIA 480
Qy 481 LEGALKLKEISYTHAEAYAGELKHGFLALIDAMPVIVVAPNNLEKLNIEEVRAR 540
Db 481 LEGALKLKEISYTHAEAYAGELKHGFLALIDAMPVIVVAPNNLEKLNIEEVRAR 540
Qy 541 GGQLYVFADQAGFVSSDNHIIEMPHVEEVIAPFYTVPLQLLAYHVALIKGTDVDDPR 600
Db 541 GGQLYVFADQAGFVSSDNHIIEMPHVEEVIAPFYTVPLQLLAYHVALIKGTDVDDPR 600
Qy 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 4
Q31UM9_SHIDS
ID Q31UM9_SHIDS PRELIMINARY; PRT; 609 AA.
AC Q31UM9;
DT 06-DEC-2005, integrated into UniProtKB/TrEMBL.
DT 06-DEC-2005, sequence version 1.
DE L-glutamine:D-fructose-6-phosphate aminotransferase.
GN Name=glms; OrderedLocusNames=SBO_3758;
OS Shigella boydii serotype 4 (strain Sb227).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=300268;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=16275786; DOI=10.1093/nar/gki954;
RA Yang F., Yang J., Zhang X., Chen L., Jiang Y., Yan Y., Tang X.,
RA Wang J., Xiong Z., Dong J., Xue Y., Zhu Y., Xu X., Sun L., Chen S.,
RA Nie H., Peng J., Xu J., Wang Y., Yuan Z., Wen Y., Yao Z., Shen Y.,
RA Qiang B., Hou Y., Yu J., Jin Q.;
RT "Genome dynamics and diversity of Shigella species, the etiologic
RT agents of bacillary dysentery."
RL Nucleic Acids Res. 33:6445-6458(2005).
CC -----
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL: CP000036; AB668229.1; -; Genomic_DNA.
DR GO: GO:0005737; C:cytoplasm; IEA.
DR GO: GO:0004360; F:glutamine-fructose-6-phosphate transaminase. . ; IEA.
DR GO: GO:0005529; F:sugar binding; IEA.
DR GO: GO:0016740; F:transferase activity; IEA.
DR GO: GO:0016051; P:carbohydrate biosynthesis; IEA.
DR GO: GO:0005975; P:carbohydrate metabolism; IEA.
DR GO: GO:0008152; P:metabolism; IEA.
KW Aminotransferase; Complete proteome; Transferase.
SQ SEQUENCE 609 AA; 66898 MW; 989B94DC79C348D4 CRC64;

Query Match 99.8%; Score 3079; DB 2; Length 609;
Best Local Similarity 99.7%; Pred. No. 7.7e-186;
Matches 607; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MCGIVGAIAORDVAEILLEGRLREYRGYSAGLAVVDAGHMTLRRLRGKVMQAAAE 60
Db 1 MCGIVGAIAORDVAEILLEGRLREYRGYSAGLAVVDAGHMTLRRLRGKVMQAAAE 60

Qy 61 EHPLHGCTGIAHTRWATHGSPSEVNAHPHVSEHIVVHNGIIEHNEHPLREBELKARGYTFV 120
```

```
Db      61  EHPLHGCGTGAHTRWATHGSPSEANAHVPHVSEHIMVWHNGIENHEPLREELKARGYTFV 120
Qy      121 SETDTEVIAHLVNWELKQGGLTLEAVLRAIPQIRGAYGTVMDSRHPDPTLLAARSGSPLV 180
Db      121 SETDTEVIAHLVNWELKQGGLTLEAVLRAIPQIRGAYGTVMDSRHPDPTLLAARSGSPLV 180
Qy      181 IGLGMENFTASDQALLPVTREPIFLEEGDIAEITRRSVNIFDKTGAEVKQDIESNLQ 240
Db      181 IGLGMENFTASDQALLPVTREPIFLEEGDIAEITRRSVNIFDKTGAEVKQDIESNLQ 240
Qy      241 YDAGDKGIYRHYMQKEIYEQPNAIKNTLTGRISHGQVDLSLGPNADELLSKVEHIQILA 300
Db      241 YDAGDKGIYRHYMQKEIYEQPNAIKNTLTGRISHGQVDLSLGPNADELLSKVEHIQILA 300
Qy      301 CGTSYNSGMVSRVWFSLAGIPCDVEIASEFRYKRSKAVRRNSLMTLSQGETADTLAGL 360
Db      301 CGTSYNSGMVSRVWFSLAGIPCDVEIASEFRYKRSKAVRRNSLMTLSQGETADTLAGL 360
Qy      361 RLSKELGYLGLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQTLVLLMLVAKL 420
Db      361 RLSKELGYLGLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQTLVLLMLVAKL 420
Qy      421 SRLKGLDASTEHDIHVHGLQALPSRIEQMLSQDKRIEALAEFSDKXHALFLGRGDQYPTA 480
Db      421 SRLKGLDASTEHDIHVHGLQALPSRIEQMLSQDKRIEALAEFSDKXHALFLGRGDQYPTA 480
Qy      481 LEGALKLKEISYTHAAYAGELKHGPLALIDAMPVIVVAPNNELLEKLKNIEBVRAR 540
Db      481 LEGALKLKEISYTHAAYAGELKHGPLALIDAMPVIVVAPNNELLEKLKNIEBVRAR 540
Qy      541 GGOLYVFAODAGFVSSDNNHIIEMPHREVTAPIFTVPLQALLVHVVALIKGTVDQDPR 600
Db      541 GGOLYVFAODAGFVSSDNNHIIEMPHREVTAPIFTVPLQALLVHVVALIKGTVDQDPR 600
Qy      601 NLAKSVTVE 609
Db      601 NLAKSVTVE 609

RESULT 5
GLMS_ECO57
ID      GLMS_ECO57      STANDARD;      PRT;      608 AA.
AC      QBXEG2;
DT      11-JUL-2002, integrated into UniProtKB/Swiss-Prot.
DT      11-JUL-2002, sequence version 2.
DT      07-MAR-2006, entry version 24.
DE      Glucosamine--fructose-6-phosphate aminotransferase [isomerizing]
DE      (EC 2.6.1.16) (Hexosephosphate aminotransferase) (D-fructose-6-
DE      phosphate amidotransferase) (GPAT) (L-glutamine-D-fructose-6-phosphate
DE      amidotransferase) (Glucosamine-6-phosphate synthase).
GN      Name=glms; OrderedLocusNames=z5227, EC64671;
OS      Escherichia coli O157:H7.
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC      Enterobacteriaceae; Escherichia.
OX      NCBI_TaxID=83334;
RN      [1]
RN      NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP      STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
RX      MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
RA      Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA      Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA      Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA      Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Pocamousis K.,
RA      Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA      Welch R.A., Blattner F.R.;
RT      "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL      Nature 409:529-533 (2001).
RN      [2]
RN      NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP      STRAIN=O157:H7 / Sakai / RIMD 0509952 / EHEC;
RX      MEDLINE=21156231; PubMed=11258796; DOI=10.1093/dnares/8.1.11;
RA      Hayaashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
```

```
RA      Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA      Iida T., Takami H., Honda T., Saekawa C., Ogawara N., Yasunaga T.,
RA      Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT      "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT      O157:H7 and genomic comparison with a laboratory strain K-12.";
RL      DNA Res. 8:11-22(2001).
CC      !- FUNCTION: Catalyzes the first step in hexosamine metabolism,
CC      converting fructose-6P into glucosamine-6P using glutamine as a
CC      nitrogen source (By similarity).
CC      !- CATALYTIC ACTIVITY: L-glutamine + D-fructose 6-phosphate = L-
CC      glutamate + D-glucosamine 6-phosphate.
CC      !- SUBUNIT: Homodimer (By similarity).
CC      !- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC      !- SIMILARITY: In the C-terminal section; belongs to the SIS family.
CC      GPAT subfamily.
CC      !- SIMILARITY: Contains 1 type-2 glutamine amidotransferase domain.
CC      -----
CC      Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC      Distributed under the Creative Commons Attribution-NoDerivs License
CC      -----
CC      EMBL; AE005174; AAG58932.1; -; Genomic DNA.
DR      EMBL; BA000007; BAB38094.1; -; Genomic DNA.
DR      PIR; G91212; G91212.
DR      PIR; H86058; H86058.
DR      HSP; P17169; 1JXA.
DR      SMR; QBXEG2; 1-239, 243-608.
DR      GenomeReviews; BA000007_GR; EC64671.
DR      GenomeReviews; AS005174_GR; Z5227.
DR      BioCyc; ECOL8334-1:EC64671-MONOMER; -.
DR      HAMAP; MF_00164; -; 1.
DR      InterPro; IPR000583; GATase_2.
DR      InterPro; IPR005855; Glms_trans.
DR      InterPro; IPR001347; SIS.
DR      Pfam; PF00310; GATase_2; 1.
DR      Pfam; PF01380; SIS; 2.
DR      TIGRFAMs; TIGR01135; glms; 1.
DR      PROSITE; PS00443; GATASE_TYPE_II; 1.
KW      Aminotransferase; Complete proteome; Glutamine amidotransferase;
KW      Transferase.
FT      INIT_MET      0      0      By similarity.
FT      CHAIN      1      608      Glucosamine--fructose-6-phosphate
FT                                     aminotransferase [isomerizing].
FT                                     /FTid=PRO_0000135330.
FT      REGION      1      240      Glutamine amidotransferase.
FT      ACT_SITE      1      1      GATase (By similarity).
FT      ACT_SITE      603      603      Isomerization Fru-6P (By similarity).
FT      CONFLICT      566      566      H -> N (in Ref. 2).
SQ      SEQUENCE      608 AA; 66777 MW; 882EDA38B7F67148 CRC64;
```

Query Match Best Local Similarity 99.7%; Score 3077; DB 1; Length 608;

Matches 607; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CGIVGATAQRDVAEILLEGRLRLEYGYDSAGLAVVDAEGHMTLRRLRGKQVMLAQAABE 61

Db 1 CGIVGATAQRDVAEILLEGRLRLEYGYDSAGLAVVDAEGHMTLRRLRGKQVMLAQAABE 60

Qy 62 HPLHGGTGCIATRWATHGEPSEVNAHPVSHIVVWHNGIENHEPLREELKARGYTFVS 121

Db 61 HPLHGGTGCIATRWATHGEPSEVNAHPVSEHIVVWHNGIENHEPLREELKARGYTFVS 120

Qy 122 ETDTEVIAHLVNWELKQGGLTLEAVLRAIPQIRGAYGTVMDSRHPDPTLLAARSGSPLV 181

Db 121 ETDTEVIAHLVNWELKQGGLTLEAVLRAIPQIRGAYGTVMDSRHPDPTLLAARSGSPLV 180

Qy 182 GLGMENFIASDQALLPVTREPIFLEEGDIAEITRRSVNIFDKTGAEVKQDIESNLQ 241

Db 181 GLGMENFIASDQALLPVTREPIFLEEGDIAEITRRSVNIFDKTGAEVKQDIESNLQ 240

Qy 242 DAGDKGIYRHYMQKEIYEQPNAIKNTLTGRISHGQVDLSLGPNADELLSKVEHIQILAC 301

Db 241 DAGDKGIYRHYMQKEIYEQPNAIKNTLTGRISHGQVDLSLGPNADELLSKVEHIQILAC 300

```
Qy 302 GTSYNSGMVSRYPFESLAGIPCDVEIASFPRYKSAVRNLSMLTTSQSGETADTLA 361
Db 301 GTSYNSGMVSRYPFESLAGIPCDVEIASFPRYKSAVRNLSMLTTSQSGETADTLA 360
Qy 362 LSKELGYLGLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFITTLVLLMLVAKLS 421
Db 361 LSKELGYLGLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFITTLVLLMLVAKLS 420
Qy 422 RLKGLDASIEHDIHVHGLQALPSRIEOMLSQDKRIEALAEAFSDKHAFILGRGDQYPIAL 481
Db 421 RLKGLDASIEHDIHVHGLQALPSRIEOMLSQDKRIEALAEAFSDKHAFILGRGDQYPIAL 480
Qy 482 EGALKLKEISYTHAEYAAGELKHGPLALIDAMPVIVVAPNNELLEKLKSNIEEVRARG 541
Db 481 EGALKLKEISYTHAEYAAGELKHGPLALIDAMPVIVVAPNNELLEKLKSNIEEVRARG 540
Qy 542 GOLYVFADQAGFVSSDNMHIEMPHVEEVIAPFYTPVPLQLLAYHVALLKGTDDVQPRN 601
Db 541 GOLYVFADQAGFVSSDNMHIEMPHVEEVIAPFYTPVPLQLLAYHVALLKGTDDVQPRN 600
Qy 602 LAKSVTVE 609
Db 601 LAKSVTVE 608

RESULT 6
GLMS SHIFL STANDARD; PRT; 608 AA.
AC Q83IY4; Q7B299;
DT 30-AUG-2005, integrated into UniProtKB/Swiss-Prot.
DT 30-AUG-2005, sequence version 2.
DT 07-MAR-2006, entry version 22.
DE Glucosamine--fructose-6-phosphate aminotransferase [isomerizing]
DE (EC 2.6.1.16) (Hexosephosphate aminotransferase) (D-fructose-6-
DE phosphate amidotransferase) (GPAT) (L-glutamine-D-fructose-6-
DE phosphate aminotransferase) (Glucosamine-6-phosphate synthase).
GN Name=glms; OrderedLocusNames=SF3809, S3959;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA Jin Q., Yuan J., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157."
RL Nucleic Acids Res. 30:4432-4441(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T."
RL Infect. Immun. 71:2775-2786(2003).
CC -1- FUNCTION: Catalyzes the first step in hexosamine metabolism,
CC converting fructose-6P into glucosamine-6P using glutamine as a
CC nitrogen source (By similarity).
CC -1- CATALYTIC ACTIVITY: L-glutamine + D-fructose 6-phosphate = L-
CC glutamate + D-glucosamine 6-phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC -1- SIMILARITY: In the C-terminal section; belongs to the SIS family.
CC GPAT subfamily.
```

```
CC -1- SIMILARITY: Contains 1 type-2 glutamine amidotransferase domain.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AE005674; AAN45249.1; -; Genomic DNA.
DR EMBL; AE016991; AAP18948.1; -; Genomic DNA.
DR HSSP; P17169; 1JXA.
DR SMR; Q83IY4; 1-239, 243-608.
DR GenomeReviews; AB014073.GR; S3959.
DR GenomeReviews; AB005674.GR; SF3809.
DR BioCyc; SFLE198214; AAN45249.1-MONOMER; -.
DR HAMAP; MF 00164; -; 1.
DR InterPro; IPR000583; Glucase_2.
DR InterPro; IPR005855; Glms_tfans.
DR InterPro; IPR001347; SIS_tfans.
DR Pfam; PF00310; Glucase_2; 1.
DR Pfam; PF01380; SIS; 2.
DR TIGRFAMs; TIGR01135; glms; 1.
DR PROSITE; PS00443; GATASE_TYPE_II; 1.
KW Aminotransferase; Complete proteome; Glutamine amidotransferase;
KW Transferase.
FT INIT_MET 0 0 By similarity.
FT CHAIN 1 608 Glucosamine--fructose-6-phosphate
FT aminotransferase [isomerizing].
FT /FTID=PRO 0000135377.
FT REGION 1 240 Glutamine amidotransferase.
FT ACT_SITE 1 1 GATase (By similarity).
FT ACT_SITE 603 603 Isomerization Fru-6P (By similarity).
SQ SSQUENCE 608 AA; 66735 MW; BOCEDA38B6F00F7D CRC64;

Query Match 99.7%; Score 3077; DB 1; Length 608;
Best Local Similarity 99.8%; Pred. No. 1e-185;
Matches 607; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CGIVGAIAQRDVAEILLEGRLLERYGYDSAGLAVVDAEGHMTLRLKGVQMLAAAE 61
Db 1 CGIVGAIAQRDVAEILLEGRLLERYGYDSAGLAVVDAEGHMTLRLKGVQMLAAAE 60
Qy 62 HPLHGGTGIAHTRWATHGEPSEVNAHPHVSEHIVVHNGIIEHNEPLEBELKARGYTFVS 121
Db 61 HPLHGGTGIAHTRWATHGEPSEANAHPHVSEHIVVHNGIIEHNEPLEBELKARGYTFVS 120
Qy 122 ETDETVIAHLVNLKOGGTLREAVLRAIPOLRGAGVTIVMDSRHPDPTLLAARSGSLVI 181
Db 121 ETDETVIAHLVNLKOGGTLREAVLRAIPOLRGAGVTIVMDSRHPDPTLLAARSGSLVI 180
Qy 182 GLGMGFNTASDQALLPVTRRFIFLEEGDIAETRRSVNIFDKTGAEVKRDIESNLQY 241
Db 181 GLGMGFNTASDQALLPVTRRFIFLEEGDIAETRRSVNIFDKTGAEVKRDIESNLQY 240
Qy 242 DAGDKGIYRHYMQEYIEQPNALKNTLTGRISHGQVDLSLGNADLELLKVEHIQILAC 301
Db 241 DAGDKGIYRHYMQEYIEQPNALKNTLTGRISHGQVDLSLGNADLELLKVEHIQILAC 300
Qy 302 GTSYNSGMVSRYPFESLAGIPCDVEIASFPRYKSAVRNLSMLTTSQSGETADTLA 361
Db 301 GTSYNSGMVSRYPFESLAGIPCDVEIASFPRYKSAVRNLSMLTTSQSGETADTLA 360
Qy 362 LSKELGYLGLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFITTLVLLMLVAKLS 421
Db 361 LSKELGYLGLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFITTLVLLMLVAKLS 420
Qy 422 RLKGLDASIEHDIHVHGLQALPSRIEOMLSQDKRIEALAEAFSDKHAFILGRGDQYPIAL 481
Db 421 RLKGLDASIEHDIHVHGLQALPSRIEOMLSQDKRIEALAEAFSDKHAFILGRGDQYPIAL 480
Qy 482 EGALKLKEISYTHAEYAAGELKHGPLALIDAMPVIVVAPNNELLEKLKSNIEEVRARG 541
Db 481 EGALKLKEISYTHAEYAAGELKHGPLALIDAMPVIVVAPNNELLEKLKSNIEEVRARG 540
Qy 542 GOLYVFADQAGFVSSDNMHIEMPHVEEVIAPFYTPVPLQLLAYHVALLKGTDDVQPRN 601
Db 541 GOLYVFADQAGFVSSDNMHIEMPHVEEVIAPFYTPVPLQLLAYHVALLKGTDDVQPRN 600
```

| | | | |
|--|--|---|-----|
| Db | 541 | GGQYVFPADQAGFVSSDNMHIEMPHVEEVIAPFYTVPLQLLAYHVALIKGTVDQPRN | 600 |
| Qy | 602 | LAKSVTVE 609 | |
| Db | 601 | LAKSVTVE 608 | |
| RESULT 7 | | | |
| Q3YVN3 | SHISS | | |
| ID | Q3YVN3_SHISS | PRELIMINARY; PRT; 609 AA. | |
| AC | Q3YVN3; | | |
| DT | 27-SEP-2005, | integrated into UniProtKB/TrEMBL. | |
| DT | 27-SEP-2005, | sequence version 1. | |
| DT | 07-FEB-2006, | entry version 5. | |
| DE | L-glutamine-D-fructose-6-phosphate aminotransferase | | |
| GN | Name=glms; OrderedLocNames=SSO_3890; ORFNames=SSO_3890; | | |
| OS | Shigella sonnei (strain Ss046). | | |
| OC | Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; | | |
| OC | Enterobacteriaceae; Shigella. | | |
| OX | NCBI_TaxID=300269; | | |
| RN | [1] | | |
| RP | NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA]. | | |
| RX | PubMed=16275786; DOI=10.1093/nar/gki954; | | |
| RA | Yang F., Yang J., Zhang X., Chen L., Jiang Y., Yan Y., Tang X., | | |
| RA | Wang J., Xiong Z., Dong J., Xue Y., Zhu Y., Xu X., Sun L., Chen S., | | |
| RA | Nie H., Peng J., Xu J., Wang Y., Yuan Z., Wen Y., Yao Z., Shen Y., | | |
| RA | Qiang B., Hou Y., Yu J., Jin Q., | | |
| RT | "Genome dynamics and diversity of Shigella species, the etiologic | | |
| RT | agents of bacillary dysentery."; | | |
| RL | Nucleic Acids Res. 33:6445-6458 (2005). | | |
| CC | Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms | | |
| CC | Distributed under the Creative Commons Attribution-NoDerivs License | | |
| CC | ----- | | |
| DR | EMBL; CP0000038; AA290429.1; -; Genomic_DNA. | | |
| DR | GO; GO:0005737; C:cytoplasm; IEA. | | |
| DR | GO; GO:0004360; F:glutamine-fructose-6-phosphate transaminase. . .; IEA. | | |
| DR | GO; GO:0005529; F:sugar binding; IEA. | | |
| DR | GO; GO:0016740; F:transferase activity; IEA. | | |
| DR | GO; GO:0016051; P:carbohydrate biosynthesis; IEA. | | |
| DR | GO; GO:0005975; P:carbohydrate metabolism; IEA. | | |
| DR | GO; GO:0008152; P:metabolism; IEA. | | |
| DR | InterPro; IPR0005853; GATase.2. | | |
| DR | InterPro; IPR001347; SIS_trans. | | |
| DR | Pfam; PF00310; GATase.2; 1. | | |
| DR | Pfam; PF01380; SIS; 2. | | |
| DR | TIGRFAMs; TIGR01135; glms; 1. | | |
| DR | PROSITE; PS00443; GAVASE TYPE II; UNKNOWN 1. | | |
| KW | Aminotransferase; Complete proteome; Transferase. | | |
| SQ | SEQUENCE 609 AA; 66844 MW; 9BE75125C4D348D4 CRC64; | | |
| Query Match 99.6%; Score 3073; DB 2; Length 609; | | | |
| Best Local Similarity 99.7%; Pred. No. 1.9e-185; | | | |
| Matches 607; Conservative 0; Mismatches 2; Indels 0; Gaps 0; | | | |
| Qy | 1 | MCIGVCAIQRDVAEITLLRGLRLEVRGYDSAGLAVDVAEGHMTLRRLGKVMQLAAAE | 60 |
| Db | 1 | MCIGVCAIQRDVAEITLLRGLRLEVRGYDSAGLAVDVAEGHMTLRRLGKVMQLAAAE | 60 |
| Qy | 61 | EHPHGGTGIAHTRWATHGEPSEVNAHPVSHIVVHNGIIEHPELRLEELKARGYTFV | 120 |
| Db | 61 | EHPLDGGTGIAHTRWATHGEPSEVNAHPVSHIVVHNGIIEHPELRLEELKARGYTFV | 120 |
| Qy | 121 | SETDTEVIAHLNVNWLKQGGTLREAVLRAIPOLRGAYGVIMDSRHPDTLLAARSGSPLV | 180 |
| Db | 121 | SETDTEVIAHLNVNWLKQGGTLREAVLRAIPOLRGAYGVIMDSRHPDTLLAARSGSPLV | 180 |
| Qy | 181 | IGLGWGENFIADQALLPVRFRFLEEGDIAETIRRSVNI FDKTGAEVKQDIESNLQ | 240 |
| Db | 181 | IGLGWGENFIADQALLPVRFRFLEEGDIAETIRRSVNI FDKTGAEVKQDIESNLQ | 240 |
| Qy | 241 | YDAGDKGIYRHYMQEIEYQPNAIKNTLTGRISHGQVDLSLGNADLLSKVEHIQILA | 300 |

| | | | |
|------------|--|--|-----|
| Db | 241 | YDAGDKGIYRHYMQEIEYQPNAIKNTLTGRISHGQVDLSLGNADLLSKVEHIQILA | 300 |
| Qy | 301 | COTSYNSGWSRYWFSFESLAGIPCDVEIASFERYKSAVRNSLMITLSQSGTADTLAGL | 360 |
| Db | 301 | COTSYNSGWSRYWFSFESLAGIPCDVEIASFERYKSAVRNSLMITLSQSGTADTLAGL | 360 |
| Qy | 361 | RLSKELGYLGLSLAICNVPGSSILVRESDLALMTNAGTEIGVASTKATFTQLTVLLMLVAKL | 420 |
| Db | 361 | RLSKELGYLGLSLAICNVPGSSILVRESDLALMTNAGTEIGVASTKATFTQLTVLLMLVAKL | 420 |
| Qy | 421 | SRKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEFSDXKHALFLGRGQYPIA | 480 |
| Db | 421 | SRKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEFSDXKHALFLGRGQYPIA | 480 |
| Qy | 481 | LEGALKLKEISYIHAEAYAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEVRAR | 540 |
| Db | 481 | LEGALKLKEISYIHAEAYAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEVRAR | 540 |
| Qy | 541 | GGQLYVFPADQAGFVSSDNMHIEMPHVEEVIAPFYTVPLQLLAYHVALIKGTVDQPR | 600 |
| Db | 541 | GGQLYVFPADQAGFVSSDNMHIEMPHVEEVIAPFYTVPLQLLAYHVALIKGTVDQPR | 600 |
| Qy | 601 | NLAKSVTVE 609 | |
| Db | 601 | NLAKSVTVE 609 | |
| RESULT 8 | | | |
| GLMS_ECOL6 | | | |
| ID | GLMS_ECOL6 | STANDARD; PRT; 608 AA. | |
| AC | Q8PBT4; | | |
| DT | 06-JUN-2003, | integrated into UniProtKB/Swiss-Prot. | |
| DT | 06-JUN-2003, | sequence version 2. | |
| DT | 07-MAR-2006, | entry version 17. | |
| DE | Glucosamine--fructose-6-phosphate aminotransferase [isomerizing] | | |
| DE | (EC 2.6.1.16) (Hexosephosphate aminotransferase) (D-fructose-6- | | |
| DE | phosphate amidotransferase) (GFAT) (L-glutamine-D-fructose-6-phosphate | | |
| DE | amidotransferase) (Glucosamine-6-phosphate synthase). | | |
| GN | Name=glms; OrderedLocNames=c4654; | | |
| OS | Escherichia coli O6. | | |
| OC | Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; | | |
| OC | Enterobacteriaceae; Escherichia. | | |
| OX | NCBI_TaxID=217992; | | |
| RN | [1] | | |
| RP | NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA]. | | |
| RC | STRAIN=O6:H1 / CFT073 / ATCC 700928 / UPEC; | | |
| RX | MEDLINE=22388234; PubMed=12471157; DOI=10.1073/pnas.252529799; | | |
| RA | Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P., | | |
| RA | Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D., | | |
| RA | Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T., | | |
| RA | Mobley H.L.T., Donnenberg M.S., Blattner F.R.; | | |
| RT | "Extensive mosaic structure revealed by the complete genome sequence | | |
| RT | of uropathogenic Escherichia coli."; | | |
| RL | Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002). | | |
| CC | !- FUNCTION: Catalyzes the first step in hexosamine metabolism, as a | | |
| CC | converting fructose-6P into glucosamine-6P using glutamine as a | | |
| CC | nitrogen source (By similarity). | | |
| CC | !- CATALYTIC ACTIVITY: L-glutamine + D-fructose 6-phosphate = L- | | |
| CC | glutamate + D-glucosamine 6-phosphate. | | |
| CC | !- SUBUNIT: Homodimer (By similarity). | | |
| CC | !- SUBCELLULAR LOCATION: Cytoplasm (By similarity). | | |
| CC | !- SIMILARITY: In the C-terminal section; belongs to the SIS family. | | |
| CC | GFAT subfamily. | | |
| CC | !- SIMILARITY: Contains 1 type-2 glutamine amidotransferase domain. | | |
| CC | ----- | | |
| CC | Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms | | |
| CC | Distributed under the Creative Commons Attribution-NoDerivs License | | |
| CC | ----- | | |
| DR | EMBL; AE016769; AA083086.1; -; Genomic_DNA. | | |
| DR | HSSP; P17169; IMOS. | | |
| DR | SMR; Q8PBT4; 1-239, 243-608. | | |
| DR | GenomeReviews; AB014075_GR; c4654. | | |


```
DR Biocyc; EC01199310:C4654-MONOMER; -.
DR HAMAP; MF 00164; -. 1.
DR InterPro; IPR000583; GATase.2.
DR InterPro; IPR005855; GlnS_trans.
DR InterPro; IPR001347; SIS.
DR Pfam; PF00310; GATase_2; 1.
DR Pfam; PF01380; SIS; 2.
DR TIGRFAMs; TIGR01135; glnS; 1.
DR PROSITE; PS00443; GATASE TYPE II; 1.
KW Aminotransferase; Complete proteome; Glutamine amidotransferase;
KW Transferase.
FT INIT_MET 0 0 By similarity.
FT CHAIN 1 608 Glucosamine--fructose-6-phosphate
FT aminotransferase [isomerizing].
FT FTID=PRO 0000135329.
FT REGION 1 240 Glutamine amidotransferase.
FT ACT_SITE 1 1 Glutamine (By similarity).
FT ACT_SITE 603 603 Isomerization Fru-6P (By similarity).
SQ SEQUENCE 608 AA; 66647 MW; 44DADB5D072C65D7 CRC64;

Query Match 99.4%; Score 3067; DB 1; Length 608;
Best Local Similarity 99.5%; Pred. No. 4.4e-185;
Matches 605; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CGIVGAIAQRDVAEILLEGRLRLRYGYDSAGLAVVDAEGHMTLRLRLGKQVQLAAAE 61
DB 1 CGIVGAIAQRDVAEILLEGRLRLRYGYDSAGLAVVDAEGHMTLRLRLGKQVQLAAAE 60
QY 62 HPLHGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIENHPELRLKARGYTFVS 121
DB 61 HPLHGGTGIAHTRWATHGEPSEANAPHVSEHIVVHNGIENHPELRLKARGYTFVS 120
QY 122 EDTTEVIAHLVNWELKOGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLLARSGSPYVI 181
DB 121 EDTTEVIAHLVNWELKOGGTLREAVLRAIPOLRGAYGTVMDSRHPDTLLARSGSPYVI 180
QY 182 GLGWMGFASDQALLPVTRRIFLEEGDIAETRRSVNIFDKTGAEVKRDQIESNLQY 241
DB 181 GLGWMGFASDQALLPVTRRIFLEEGDIAETRRSVNIFDKTGAEVKRDQIESNLQY 240
QY 242 DAGDKGIYRHYMQKEIYEOPNAIKNTLTGRISHGQVDLSGLGNADLLSKVEHIQILAC 301
DB 241 DAGDKGIYRHYMQKEIYEOPNAIKNTLTGRISHGQVDLSGLGNADLLSKVEHIQILAC 300
QY 302 GTSYNGMSRYWFESLAGIPCDVEIASEFRYKSAVRNSLMTLSQSGETADTLAQLR 361
DB 301 GTSYNGMSRYWFESLAGIPCDVEIASEFRYKSAVRNSLMTLSQSGETADTLAQLR 360
QY 362 LSKELGYLGLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFITQTLTLLMLVAKLS 421
DB 361 LSKELGYLGLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFITQTLTLLMLVAKLS 420
QY 422 RLKGLDASIEHDIVHGLQALPSRIEQMLSQDKRIEALAEEDFSKHHALFLGRGDQYPIAL 481
DB 421 RLKGLDASIEHDIVHGLQALPSRIEQMLSQDKRIEALAEEDFSKHHALFLGRGDQYPIAL 480
QY 482 EGALKLKEISYTHAEYAAAGELKHGPLALIDAMPVIVVAPNNELLEKLKSNIEEVRAG 541
DB 481 EGALKLKEISYTHAEYAAAGELKHGPLALIDAMPVIVVAPNNELLEKLKSNIEEVRAG 540
QY 542 GOLYVFADQDAGFVSDNNHIIEMPHVEEVIAPFVTVPLQLLAYHVALIKGTDVDPQRN 601
DB 541 GOLYVFADQDAGFVSDNNHIIEMPHVEEVIAPFVTVPLQLLAYHVALIKGTDVDPQRN 600
QY 602 LAKSVTVE 609
DB 601 LAKSVTVE 608

RESULT 9
GLMS_SALPA STANDARD; PRT; 608 AA.
ID _GLMS_SALPA
AC Q5PKV9;
```

```
DT 30-AUG-2005, integrated into UniProtKB/Swiss-Prot.
DT 30-AUG-2005, sequence version 2.
DT 07-MAR-2006, entry version 12.
DE Glucosamine--fructose-6-phosphate aminotransferase [isomerizing]
DE (EC 2.6.1.16) (Hexosephosphate aminotransferase) (D-fructose-6-
DE phosphate amidotransferase) (GFAT) (L-glutamine-D-fructose-6-phosphate
DE amidotransferase) (Glucosamine-6-phosphate synthase).
DE Names: glnS; OrderedLocusNames=SPA3700;
OS Salmonella paratyphi-a.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=54388;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 9150 / SARB42;
RX PubMed=15531882; DOI=10.1038/ngl470;
RA McClelland M., Sanderson K.E., Clifton S.W., Latreille P.,
RA Porwollik S., Sabo A., Meyer R., Bieri T., Ozerky P., McLellan M.,
RA Harkins C.R., Wang C., Nguyen C., Berghoff A., Elliott G.,
RA Kohlberg S., Strong C., Du F., Carter J., Kremizki C., Layman D.,
RA Leonard S., Sun H., Fulton L., Nash W., Miner T., Minx P., Florea L.,
RA Delehaanty K., Fronick C., Magrini V., Nhan M., Warren W., Florea L.,
RA Sleeth J., Wilson R.K.;
RT "Comparison of genome degradation in Paratyphi A and Typhi, human-
RT restricted serovars of Salmonella enterica that cause typhoid.";
RL Nat. Genet. 36:1268-1274(2004).
CC -!- FUNCTION: Catalyzes the first step in hexosamine metabolism,
CC converting fructose-6P into glucosamine-6P using glutamine as a
CC nitrogen source (By similarity).
CC -!- CATALYTIC ACTIVITY: L-glutamine + D-fructose 6-phosphate = L-
CC glutamate + D-glucosamine 6-phosphate.
CC -!- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC -!- SIMILARITY: In the C-terminal section; belongs to the SIS family.
CC -!- GPAT subfamily.
CC -!- SIMILARITY: Contains 1 type-2 glutamine amidotransferase domain.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; CP000026; AAV79492.1; -; Genomic_DNA.
DR SMR; Q5PKV9; 1-239, 243-608.
DR GenomeReviews; CP000026_GR; SPA3700.
DR HAMAP; MF 00164; -. 1.
DR InterPro; IPR000583; GATase.2.
DR InterPro; IPR005855; GlnS_trans.
DR InterPro; IPR001347; SIS.
DR Pfam; PF00310; GATase_2; 1.
DR Pfam; PF01380; SIS; 2.
DR TIGRFAMs; TIGR01135; glnS; 1.
DR PROSITE; PS00443; GATASE TYPE II; 1.
KW Aminotransferase; Complete proteome; Glutamine amidotransferase;
KW Transferase.
FT INIT_MET 0 0 By similarity.
FT CHAIN 1 608 Glucosamine--fructose-6-phosphate
FT aminotransferase [isomerizing].
FT FTID=PRO 0000135373.
FT REGION 1 240 Glutamine amidotransferase.
FT ACT_SITE 1 1 Glutamine (By similarity).
FT ACT_SITE 603 603 Isomerization Fru-6P (By similarity).
SQ SEQUENCE 608 AA; 66719 MW; EE20C7F1996B742B CRC64;

Query Match 99.1%; Score 3058; DB 1; Length 608;
Best Local Similarity 99.2%; Pred. No. 1.6e-184;
Matches 603; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 CGIVGAIAQRDVAEILLEGRLRLRYGYDSAGLAVVDAEGHMTLRLRLGKQVQLAAAE 61
DB 1 CGIVGAIAQRDVAEILLEGRLRLRYGYDSAGLAVVDAEGHMTLRLRLGKQVQLAAAE 60
QY 62 HPLHGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIENHPELRLKARGYTFVS 121
DB 61 HPLHGGTGIAHTRWATHGEPSEANAPHVSEHIVVHNGIENHPELRLKARGYTFVS 120
```

```
QY 122 ETDTVEIAHLVNWELKOGGTLREAVLRAIPQLRGAYGTVMDSRHPDPTLLAARSGSPLYI 181
DB 121 ETDTVEIAHLVNWELKOGGTLREAVLRAIPQLRGAYGTVMDSRHPDPTLLAARSGSPLYI 180
QY 182 GLGMGENFIASDQALLPVTRRFIFLEEGDIAETIRSVNIPFRTGAEVKRDIESNLQY 241
DB 181 GLGMGENFIASDQALLPVTRRFIFLEEGDIAETIRSVNIPFRTGAEVKRDIESNLQY 240
QY 242 DAGDKGIYRHYMQKEIYEQNAIKNTLTGRISHGQVDLSLGNADDELLSKVHEIQIILAC 301
DB 241 DAGDKGIYRHYMQKEIYEQNAIKNTLTGRISHGQVDLSLGNADDELLSKVHEIQIILAC 300
QY 302 GTSYNSGMVSRYPFESLAGIPCDVEIASPEYRKSAVRNLSMITLSQSGETADTLAAGR 361
DB 301 GTSYNSGMVSRYPFESLAGIPCDVEIASPEYRKSAVRNLSMITLSQSGETADTLAAGR 360
QY 362 LSKELGYLGLSLAICNVPSSSLVRESDLALMTNAGTEIGVASTKFTTQTLTVLLMLVAKLS 421
DB 361 LSKELGYLGLSLAICNVPSSSLVRESDLALMTNAGTEIGVASTKFTTQTLTVLLMLVAKLS 420
QY 422 RLKGLDASIEHDIHVHGLQALPSRIEQMLSQDKRIEALAEEDFSKHHALFLGRGDQYPIAL 481
DB 421 RLKGLDASIEHDIHVHGLQALPSRIEQMLSQDKRIEALAEEDFSKHHALFLGRGDQYPIAL 480
QY 482 EGALKLKEISYIHAEAYAAAGELKHGPLALIDAMPVIVVAPNNLEKLSNIEEVRARG 541
DB 481 EGALKLKEISYIHAEAYAAAGELKHGPLALIDAMPVIVVAPNNLEKLSNIEEVRARG 540
QY 542 GQLYVFADQAGFVSSDNMHIEMPHVEEVIAPFVTVPLQLLAYHVALIKGTDVDDQPRN 601
DB 541 GQLYVFADQAGFVSSDNMHIEMPHVEEVIAPFVTVPLQLLAYHVALIKGTDVDDQPRN 600
QY 602 LAKSVTVE 609
DB 601 LAKSVTVE 608

RESULT 10
Q57HY2_SALCH
ID Q57HY2_SALCH PRELIMINARY; PRT; 609 AA.
AC Q57HY2;
AT 10-MAY-2005, integrated into UniProtKB/TrEMBL.
DT 10-MAY-2005, sequence version 1.
DT 07-FEB-2006, entry version 4.
DE L-glutamine-D-fructose-6-phosphate aminotransferase.
GN Name=glms; OrderedLocusNames=SC3774; ORFNames=SC3774;
OS Salmonella choleraesuis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=591;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=SC-867;
RX PubMed=15781495; DOI=10.1093/nar/gki297;
RA Chiu C.-H., Tang P., Chu C., Hu S., Bao Q., Yu J., Chou Y.-Y.,
RA Wang H.-S., Lee Y.-S.;
RT "The genome sequence of Salmonella enterica serovar Choleraesuis, a
RL highly invasive and resistant zoonotic pathogen.";
RL Nucleic Acids Res. 33:1690-1698(2005).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; AS017220; AAX67680.1; -; Genomic_DNA.
DR SNR; Q57HY2; 2-240, 244-609.
DR GO; GO:0005737; C:cycloplasm; IEA.
DR GO; GO:0004360; F:glutamine-fructose-6-phosphate transaminase. . .; IEA.
DR GO; GO:0005529; F:sugar binding; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0016051; P:carbohydrate biosynthesis; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR000583; GATase_2.
```

```
DR InterPro; IPR005955; Glms_trans.
DR InterPro; IPR001347; SIS.
DR Pfam; PF00310; GATase_2; 1.
DR Pfam; PF01380; SIS; 2.
DR TIGRFAMs; TIGR01135; glms; 1.
DR PROSITE; PS00443; GATASE TYPE II; UNKNOWN 1.
KW Aminotransferase; Complete proteome; Transferase.
SQ SEQUENCE 609 AA; 66849 MW; 66849 MW; ELDLFF6FIAP35FSC CRC64;

Query Match
Best Local Similarity 99.1%; Score 3058; DB 2; Length 609;
Matches 602; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 MCGIVGAIQORDVABILLLEGRLRLEYRGYDSAGLAVDAEGHMTLRRLGKQVQLAAQAE 60
DB 1 MCGIVGAIQORDVABILLLEGRLRLEYRGYDSAGLAVDAEGHMTLRRLGKQVQLAAQAE 60
QY 61 EHPHGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHPELREELKARGYTFV 120
DB 61 EHPHGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHPELREELKARGYTFV 120
QY 121 SETDTVEIAHLVNWELKOGGTLREAVLRAIPQLRGAYGTVMDSRHPDPTLLAARSGSPLY 180
DB 121 SETDTVEIAHLVNWELKOGGTLREAVLRAIPQLRGAYGTVMDSRHPDPTLLAARSGSPLY 180
QY 181 IGLGMGENFIASDQALLPVTRRFIFLEEGDIAETIRSVNIPFRTGAEVKRDIESNLQ 240
DB 181 IGLGMGENFIASDQALLPVTRRFIFLEEGDIAETIRSVNIPFRTGAEVKRDIESNLQ 240
QY 241 YDAGDKGIYRHYMQKEIYEQNAIKNTLTGRISHGQVDLSLGNADDELLSKVHEIQIILA 300
DB 241 YDAGDKGIYRHYMQKEIYEQNAIKNTLTGRISHGQVDLSLGNADDELLSKVHEIQIILA 300
QY 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASPEYRKSAVRNLSMITLSQSGETADTLAAGL 360
DB 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASPEYRKSAVRNLSMITLSQSGETADTLAAGL 360
QY 361 RLSKELGYLGLSLAICNVPSSSLVRESDLALMTNAGTEIGVASTKFTTQTLTVLLMLVAKL 420
DB 361 RLSKELGYLGLSLAICNVPSSSLVRESDLALMTNAGTEIGVASTKFTTQTLTVLLMLVAKL 420
QY 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMLSQDKRIEALAEEDFSKHHALFLGRGDQYPIA 480
DB 421 ARLKGLDTSIEHDIHVHGLQALPSRIEQMLSQDKRIEALAEEDFSKHHALFLGRGDQYPIA 480
QY 481 LEGALKLKEISYIHAEAYAAAGELKHGPLALIDAMPVIVVAPNNLEKLSNIEEVRARG 540
DB 481 LEGALKLKEISYIHAEAYAAAGELKHGPLALIDAMPVIVVAPNNLEKLSNIEEVRARG 540
QY 541 GGQLYVFADQAGFVSSDNMHIEMPHVEEVIAPFVTVPLQLLAYHVALIKGTDVDDQPR 600
DB 541 GGQLYVFADQAGFVSSDNMHIEMPHVEEVIAPFVTVPLQLLAYHVALIKGTDVDDQPR 600
QY 601 NLAKSVTVE 609
DB 601 NLAKSVTVE 609

RESULT 11
GLMS_SALTY
ID GLMS_SALTY STANDARD; PRT; 608 AA.
AC Q8ZKX1;
AT 11-JUL-2002, integrated into UniProtKB/Swiss-Prot.
DT 11-JUL-2002, sequence version 2.
DT 07-MAR-2006, entry version 21.
DE Glucosamine-fructose-6-phosphate aminotransferase [isomerizing]
DE (EC 2.6.1.16) (Hexosephosphate aminotransferase) (D-fructose-6-
DE phosphate amidotransferase) (GFAT) (L-glutamine-D-fructose-6-phosphate
DE amidotransferase) (Glucosamine-6-phosphate synthase).
GN Name=glms; OrderedLocusNames=STM3861;
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
```

```

OX NCBI_TaxID=602;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=L72 / SGSC1412 / ATCC 700720; DOI=10.1038/35101614;
RA MEDLINE=21534948; PubMed=11677608; DOI=10.1038/35101614;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT L72.";
RL Nature 413:852-856(2001).
CC -1- FUNCTION: Catalyzes the first step in hexosamine metabolism,
CC converting fructose-6P into glucosamine-6P using glutamine as a
CC nitrogen source (By similarity).
CC -1- CATALYTIC ACTIVITY: L-glutamine + D-fructose 6-phosphate = L-
CC glutamate + D-glucosamine 6-phosphate.
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC -1- SIMILARITY: In the C-terminal section; belongs to the SIS family.
CC GPAT subfamily.
CC -1- SIMILARITY: Contains 1 type-2 glutamine amidotransferase domain.
CC -----
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR ENBL; AE008880; AAU22719.1; -; Genomic_DNA.
DR HSSP; P17169; IJXA.
DR SMR; Q82KX1; 1-239, 243-608.
DR GenomeReviews; AE006468 GR; STM3861.
DR StyGene; SG7777?; glms.
DR BioCyc; STY99287:STM3861-MONOMER; -.
DR HAMAP; MF_00164; -; 1.
DR InterPro; IPR000583; GATase_2.
DR InterPro; IPR005855; Glms_trans.
DR InterPro; IPR001347; SIS.
DR Pfam; PF00310; GATase_2; 1.
DR Pfam; PF01380; SIS; 2.
DR TIGRFAMs; TIGR01135; glms; 1.
DR PROSITE; PS00443; GATASE_TYPE_II; 1.
KW Amidotransferase; Complete proteome; Glutamine amidotransferase;
KW Transferase.
FT INIT MET 0 0 By similarity.
FT CHAIN 1 608 Glucosamine--fructose-6-phosphate
FT /FruD=PRO.0000135375.
FT REGION 1 240 Glutamine amidotransferase.
FT ACT_SITE 1 1 GATase (By similarity).
FT ACT_SITE 603 603 Isomerization Fru-6P (By similarity).
SQ SEQUENCE 608 AA; 66746 MW; C320C7F18DEDAF52 CRC64;

Query Match 99.0%; Score 3055; DB 1; Length 608;
Best Local Similarity 99.0%; Pred. No. 2.5e-184;
Matches 602; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 CGIVGATAQRDVAEILLEGRLRLRYRGYDSAGLAVDAEGHMTLRRLGKQVQLAQAAEE 61
DB 1 CGIVGATAQRDVAEILLEGRLRLRYRGYDSAGLAVDAEGHMTLRRLGKQVQLAQAAEE 60
QY 62 HPLHGGTGIAHTRWATGEPSEVNAHPVSEHIVVHNGIIEHNEPLREELKARGYTFVS 121
DB 61 HPLHGGTGIAHTRWATGEPSEVNAHPVSEHIVVHNGIIEHNEPLREELKARGYTFVS 120
QY 122 ETDTEVIAHLVNNELKQGGTLREAVLRAIPQLRGAGYGTIVMDSRHPDPTLLAARSGSLVI 181
DB 121 ETDTEVIAHLVNNELKQGGTLREAVLRAIPQLRGAGYGTIVMDTRHPDPTLLAARSGSLVI 180
QY 182 GLGMGENFIASDQALLPVTTRRFIFLBEQDIAETTRSVNIPDKTGAEVKRDIESNLQY 241
DB 181 GLGMGENFIASDQALLPVTTRRFIFLBEQDIAETTRSVNIPDKTGAEVKRDIESNLQY 240
QY 242 DAGDKGIYRHYMQEIIYEQNAIKNTLTGRISHQGVLDSELGPNADLLSKVEHIQILAC 301

```

```

DB 241 DAGDKGIYRHYMQEIIYEQNAIKNTLTGRISHQGVLDSELGPNADLLSKVEHIQILAC 300
QY 302 GTSYNSGMWSRYWFSFESLAGIPCDVEIASERFYSKSAVRNSLMITLSQSGETADTLAQLR 361
DB 301 GTSYNSGMWSRYWFSFESLAGIPCDVEIASERFYSKSAVRNSLMITLSQSGETADTLAQLR 360
QY 362 LSKELGYLGSIALCNVPGSSLVRESDLALMTNAGTEIGVASTKRAFTTQLTVLLMLVAKLS 421
DB 361 LSKELGYLGSIALCNVPGSSLVRESDLALMTNAGTEIGVASTKRAFTTQLTVLLMLVAKLS 420
QY 422 RLKGLDASIEHDIHVGLQALPSRIEQMLSQDKRIEALAEFSDKHAFIFLGRGQOYPAL 481
DB 421 RLKGLDASIEHDIHVGLQALPSRIEQMLSQDKRIEALAEFSDKHAFIFLGRGQOYPAL 480
QY 482 EGALKLKEISYIHAAYAGELKHGPLALIDAMPVIVVAPNNELLKLSNIEEVRARG 541
DB 481 EGALKLKEISYIHAAYAGELKHGPLALIDAMPVIVVAPNNELLKLSNIEEVRARG 540
QY 542 GOLYVFADQDAGFVSSDNMHIIEMPHVEEVIAPITFYTPVQLLAYHVALIKGTDVDDPRN 601
DB 541 GOLYVFADQDAGFVSSDNMHIIEMPHVEEVIAPITFYTPVQLLAYHVALIKGTDVDDPRN 600
QY 602 LAKSVTVE 609
DB 601 LAKSVTVE 608

RESULT 12
GLMS_SALTI
ID _GLMS_SALTI STANDARD; PRT; 608 AA.
AC Q8Z2Q2;
DT 11-JUL-2002, integrated into UniProtKB/Swiss-Prot.
DT 11-JUL-2002, sequence version 2.
DE Glucosamine--fructose-6-phosphate aminotransferase [isomerizing]
DE (EC 2.6.1.16) (Hexosephosphate aminotransferase) (D-fructose-6-
DE phosphate amidotransferase) (GPAT) (L-glutamine-D-fructose-6-phosphate
DE amidotransferase) (Glucosamine-6-phosphate synthase).
DE Name=gims; OrderedLocusNames=STY3917, t3658;
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CT18;
RA MEDLINE=21534947; PubMed=11677608; DOI=10.1038/35101607;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrall B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar typhi CT18.";
RL Nature 413:848-852(2001).
[2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RX DOI=10.1128/JB.185.7.2330-2337.2003;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT and CT18";
RL J. Bacteriol. 185:2330-2337(2003).
CC -1- FUNCTION: Catalyzes the first step in hexosamine metabolism,
CC converting fructose-6P into glucosamine-6P using glutamine as a
CC nitrogen source (By similarity).
CC -1- CATALYTIC ACTIVITY: L-glutamine + D-fructose 6-phosphate = L-

```


| | | | | |
|--|--|--|----------|---------------------------------------|
| FT | REGION | 1 | 241 | Glutamine amidotransferase. |
| FT | ACT_SITE | 1 | 1 | Gatase (By similarity). |
| FT | ACT_SITE | 604 | 604 | Isomerization Fru-6P (By similarity). |
| SQ | SEQUENCE | 609 AA; | 6686 MW; | BFE71B8EDFD23B5D CRC64; |
| Query Match | | | | |
| Best Local Similarity 85.1%; Pred. No. 1e-160; | | | | |
| Matches 518; Conservative 46; Mismatches 44; Indels 1; Gaps 1; | | | | |
| Qy | 2 | CGIVGAQDVAEILLEGRLRYCYGDSAGLAVDAEGHMTLRRLGKVMQLAQAAE | 61 | |
| Db | 1 | CGIVGAQDVAEILLEGRLRYCYGDSAGLAVDSEGHVARLRRLGKQVLSQAAE | 60 | |
| Qy | 62 | HLPHGGTGIAHTRWATGEPSEVNAHPVSEHIVVHNGIIEHPEELRKARGTYFVS | 121 | |
| Db | 61 | HELHGGTGIAHTRWATGEPSEENAHPHISEHITIVHNGIIEHPELRMIGRGYFVS | 120 | |
| Qy | 122 | ETDTEVIAHLVNNELKQ-GGTLREAVLRAIPQLRGAYGTVMDSRHPDTLAARSGPLV | 180 | |
| Db | 121 | ETDTEVIAHLVHFPEQKQGGTLVEVVKRVPQLRGAYGMVLDNRDSSVLVAARSGPLV | 180 | |
| Qy | 181 | IGLGMENFTASDQALLPYTRFIFLEEGDIAEITRVSNIYFDKTAEVKRODIESNLQ | 240 | |
| Db | 181 | IGRGVGENFTASDQALLPYTRRFMFLEEGDVAEITRDVRFVDFKSGQLATREEIESKVS | 240 | |
| Qy | 241 | YDAGDKGIYRHYMKEIYEQPNAIKNTLTGRISHGOVDLSELGPNADLLSKVEHIQILA | 300 | |
| Db | 241 | YDAGDKGAYRHYMKEIYEQPMAIKNTLEGRFSHGEINLSELGPKADELLAKVHVQIIA | 300 | |
| Qy | 301 | CGTSYNGMVSRYWFESLAGIPCDVEIASEFRYKRSVNRNSLMTLSQSGETADTLAQL | 360 | |
| Db | 301 | CGTSYNGMVSRYWFELAGIPCDVEIASEFRYKRPVNRNSLMTLSQSGETADTLAAL | 360 | |
| Qy | 361 | RLSKELGYLGSIAICNVPGSSLVRESLALMTWAGTEIGVASTKAFPTQLTVLLMLVAKL | 420 | |
| Db | 361 | RLSKELGYLGSIAICNVAGSSLVRESDMALMTKAGVEIGVASTKAFPTQLTVLLMLVARV | 420 | |
| Qy | 421 | SRUKGLDASTEHDIHVHGLQALPSRIEOMLSQDKRIEALAEFDDKHALFLGRGDQYPIA | 480 | |
| Db | 421 | GRLRGMDAQIEHDIHVHGLQALPARIEQMLSDKLIESLAGFDDKHALFLGRGDQYPIA | 480 | |
| Qy | 481 | LEGALKLKEISYTHAEYAAAGELKHGPLALIDAMPVIVVAPNNLEKLKSNIEEVRAR | 540 | |
| Db | 481 | MEGALKLKEISYTHAEYAAAGELKHGPLALIDAMPVIVVAPNNLEKLKSNIEEVRAR | 540 | |
| Qy | 541 | CGOLYVFADQAGFVSDNHHIEMPHVERVIAPFYVTPLOLLAHYVALIKGTVDVQPR | 600 | |
| Db | 541 | CGELYVFADEDAGFTSENKKIIPLPHIEEVIAPFYVTPLOLLSYHVALIKGTVDVQPR | 600 | |
| Qy | 601 | NLAKSVTVE 609 | | |
| Db | 601 | NLAKSVTVE 609 | | |
| RESULT 14 | | | | |
| GLMS_YERPE | | | | |
| AC | Q82958; | STANDARD; | PRT; | 608 AA. |
| DT | 11-JUL-2002, | integrated into UniProtKB/Swiss-Prot. | | |
| DT | 11-JUL-2002, | sequence version 2. | | |
| DT | 07-MAR-2006, | entry version 32. | | |
| DE | Glucoseamine--fructose-6-phosphate aminotransferase [isomerizing] | | | |
| DE | (EC 2.6.1.16) (Hexosephosphate aminotransferase) (D-fructose-6-phosphate | | | |
| DE | amidotransferase) (GATase) (L-glutamine-D-fructose-6-phosphate | | | |
| GN | Name=glms; OrderedLocusNames=YPO4118, Y4132, YP4025; | | | |
| OS | Yersinia pestis. | | | |
| OC | Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; | | | |
| OC | Enterobacteriaceae; Yersinia. | | | |
| OX | NCBI_TaxID=632; | | | |
| RN | [1] | | | |
| RP | NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA]. | | | |
| RC | STRAIN=CO-92 / Biovar Orientalis; | | | |

| | | |
|------|--|-----------------------------------|
| RX | MEDLINE=21470413; PubMed=11586360; DOI=10.1038/35097083; | |
| RA | Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G., | |
| RA | Prentice M.B., Sebaihia M., James K.D., Churcher C.M., Mungall K.L., | |
| RA | Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.-M., | |
| RA | Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G., | |
| RA | Fellwell T., Hamlin N., Holtroyd S., Jagers K., Karlyshev A.V., | |
| RA | Leatherbarrow S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.M., | |
| RA | Simmonds M., Skellon J., Stevens K., Whitehead S., Barrall B.G.; | |
| RT | "Genome sequence of Yersinia pestis, the causative agent of plague."; | |
| RL | Nature 413:523-527(2001). | |
| RN | [2] | |
| RP | NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA]. | |
| RC | STRAIN=KIMS / Biovar Mediaevalis; | |
| RX | MEDLINE=22137863; PubMed=12142430; | |
| RX | DOI=10.1128/JB.184.16.4601-4611.2002; | |
| RA | Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P., | |
| RA | Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C., Liss P., | |
| RA | Rethersson J.D., Lindler L.E., Brubaker R.R., Plano G.V., | |
| RA | Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R., | |
| RA | Perry R.D.; | |
| RT | "Genome sequence of Yersinia pestis KIM."; | |
| RL | J. Bacteriol. 184:4601-4611(2002). | |
| RN | [3] | |
| RP | NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA]. | |
| RC | STRAIN=91001 / Biovar Mediaevalis; | |
| RX | PubMed=15368893; DOI=10.1093/dnares/11.3.179; | |
| RA | Song Y., Tong Z., Wang J., Wang L., Guo Z., Han Y., Zhang J., Pei D., | |
| RA | Zhou D., Qin H., Pang X., Han Y., Zhai J., Li M., Cui B., Qi Z., | |
| RA | Jin L., Dai R., Chen F., Li S., Ye C., Du Z., Lin W., Wang J., Yu J., | |
| RA | Yang H., Wang J., Huang P., Yang R.; | |
| RT | "Complete genome sequence of Yersinia pestis strain 91001, an isolate | |
| RT | avirulent to humans."; | |
| RL | DNA Res. 11:179-197(2004). | |
| CC | -I- FUNCTION: Catalyzes the first step in hexosamine metabolism, | |
| CC | converting fructose-6P into glucosamine-6P using glutamine as a | |
| CC | nitrogen source (By similarity). | |
| CC | -I- CATALYTIC ACTIVITY: L-glutamine + D-fructose 6-phosphate = L- | |
| CC | glutamate + D-glucosamine 6-phosphate. | |
| CC | -I- SUBCELLULAR LOCATION: Cytoplasm (By similarity). | |
| CC | -I- SIMILARITY: In the C-terminal section; belongs to the SIS family. | |
| CC | GFAT subfamily. | |
| CC | -I- SIMILARITY: Contains 1 type-2 glutamine amidotransferase domain. | |
| CC | Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms | |
| CC | Distributed under the Creative Commons Attribution-NoDerivs License | |
| CC | ----- | |
| EMBL | AJ414160; CAC93567.1; -; Genomic_DNA. | |
| EMBL | AE014014; AAM87674.1; -; Genomic_DNA. | |
| EMBL | AE017142; AAS64164.1; -; Genomic_DNA. | |
| DR | PIR; AB0500; AB0500. | |
| DR | HSSP; P17169; 1MOS. | |
| DR | SNR; Q82958; 1-239, 243-608. | |
| DR | GenomeReviews; AE009952_GR; Y4132. | |
| DR | GenomeReviews; AE017042_GR; YP4025. | |
| DR | GenomeReviews; AL590842_GR; YPO4118. | |
| DR | BioCyc; YPES187410.Y4132-MONOMER; -. | |
| DR | BioCyc; YPES632.YPO4118-MONOMER; -. | |
| DR | HMAP; MF_00164; -; 1. | |
| DR | InterPro; IPR000583; GATase_2. | |
| DR | InterPro; IPR005855; Glms_trans. | |
| DR | InterPro; IPR001347; SIS. | |
| DR | Pfam; PF00310; GATase_2; 1. | |
| DR | Pfam; PF01380; SIS; 2. | |
| DR | TIGRFAMs; TIGR01135; glms; 1. | |
| DR | PROSITE; PS00443; GATASE_TYPE_II; 1. | |
| KW | Aminotransferase; Complete proteome; Glutamine amidotransferase; | |
| KW | Transferase. | |
| FT | INIT MET 0 0 | By similarity. |
| FT | CHAIN 1 608 | Glucosamine--fructose-6-phosphate |
| FT | | amidotransferase [isomerizing]. |
| FT | | /FTid=PRO_0000135418. |
| FT | REGION 1 240 | Glutamine amidotransferase. |
| FT | ACT_SITE 1 1 | GATase (By similarity). |

```
FT ACT SITE 603 603 Isomerization Fru-6P (By similarity).
SQ SEQUENCE 608 AA; 66394 MW; 1655FD0ADB16CCD6 CRC64;

Query Match
Best Local Similarity 86.7%; Score 2677; DB 1; Length 608;
Matches 517; Conservative 44; Mismatches 47; Indels 0; Gaps 0;

QY 2 CGIVGAIQAQRDVAREILLEGRLRLEGRYGYDSAGLAVVDAEGHMTLRRLGKVMQLAAAE 61
Db 1 CGIVGAVQAQRDIAEILLEGRLRLEGRYGYDSAGLAVVDSGHLTRLRRVGVKVALSDAAEK 60

QY 62 HPLHGGTGTIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEPLREELKARGTYFVS 121
Db 61 QDLHGGTGTIAHTRWATHGEPSEANAPHVSDYSIVVHNGIIEHNEPLRELLISRGYRFS 120

QY 122 ETDETEVIAHLVNWELKOGGTLRREAVRAIPQLRGAYGTVIMDSRHPDPTLLAARSGSPVLI 181
Db 121 ETDETEVIAHLVHVEQOQGSLLVVKRVIPLQLRGAYGTVVWDSRDLRLIARSGSPVLI 180

QY 182 GLGNGENFIASDQALLPVTRRRFIFLEEGDIAETRRSVNIFDKTGAEVKRODIESNLOY 241
Db 181 GCGVGENFIASDQALLPVTRRRFIFLEEGDVVEVTRRSISIFDKQNAIERPEIESQVQY 240

QY 242 DAGDKGIRYHYMQEIIYEOPNAIKNTLTGRISHGQVDLSGLGNADLLSKVEHIQILAC 301
Db 241 DAGDKGIRYHYMQEIIYEOPMAIKNTLEGLRSHGMIDLSELGPKADALLAEVQHIQIAC 300

QY 302 GTSYNSGMWSRYWPFESLAGIPCDVEIASPEFRYKSAVRNSLMITLSQSGETADTLAAGR 361
Db 301 GTSYNSGMWSRYWPFESLAGVPCDVEIASPEFRYKSAVRPNLSLITLSQSGETADTLAALR 360

QY 362 LSKELGYLGSALICNVPGSSLVRESLALMTNAGTIGVASTKFTTQTLVLLMLVAKLS 421
Db 361 LSKELGYLGSALICNVAGSSLVRESLALMTKAGTIGVASTKFTTQTLVLLMLVGRIG 420

QY 422 RLKGLDASIEHDIVHGLQALPSRIEQMLSQDKRIEALAEFDSKHALFLRGDQVPIAL 481
Db 421 KLKGDASLEHDIVHALQALPARIEQMLSUDKTIKTEALAEFSFGKHALFLRGDQVPIAM 480

QY 482 EGALKLKEISYIHAEYAAAGELKHGFLALIDAMPVIVVAPNDELLEKUKSNIEVRARG 541
Db 481 EGALKLKEISYIHAEYAAAGELKHGFLALIDAMPVIVVAPNDELLEKUKSNIEVRARG 540

QY 542 GOLYVFADODAGVSSDNNHHIEMPHVEEVIAPFVTVPLQLLAVHVALIKCTDQVDPQN 601
Db 541 GLLYVFADQDAGTDSGKKIQLPVEEIIAPFVTVPLQLLTHVALIKCTDQVDPQN 600

QY 602 LAKSVTVE 609
Db 601 LAKSVTVE 608

RESULT 15
GLMS_YERPS STANDARD; PRT; 608 AA.
AC O663R1;
DT 30-AUG-2005, integrated into UniProtKB/Swiss-Prot.
DT 30-AUG-2005, sequence version 2.
DT 30-MAR-2006, entry version 16.
DE Glucosamine-6-phosphate aminotransferase (isomerizing)
DE (EC 2.6.1.16) (Hexosephosphate aminotransferase) (D-fructose-6-
DE phosphate amidotransferase) (GPA) (L-glutamine-D-fructose-6-phosphate
DE amidotransferase) (Glucosamine-6-phosphate synthase).
GN Names:glms; OrderedLocusNames=YPTB3964;
OS Versinia pseudotuberculosis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Versinia.
OX NCBI_TaxID=633;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=IP32953 / Serotype I;
RA PubMed=15358858; DOI=10.1073/pnas.0404012013;
Chain P.S.G., Carniel E., Larimer F.W., Lamerdin J., Stoutland P.O.,
```

```
RA Regala W.M., Georgescu A.M., Vergez L.M., Land M.L., Motin V.L.,
RA Brubaker R.R., Fowler J., Hinnebusch J., Marceau M., Medigue C.,
RA Simonet M., Chenal-Francisque V., Souza B., Dacheux D., Elliott J.M.,
RA Derbise A., Hauser L.J., Garcia E.;
RT "Insights into the evolution of Versinia pseudotuberculosis":
RT comparison with Versinia pseudotuberculosis";
RL Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831(2004).
CC -!- FUNCTION: Catalyzes the first step in hexosamine metabolism,
CC converting fructose-6P into glucosamine-6P using glutamine as a
CC nitrogen source (By similarity).
CC -!- CATALYTIC ACTIVITY: L-glutamine + D-fructose 6-phosphate = L-
CC glutamate + D-glucosamine 6-phosphate.
CC -!- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC -!- SIMILARITY: In the C-terminal section; belongs to the SIS family.
CC GFAT subfamily.
CC -!- SIMILARITY: Contains 1 type-2 glutamine amidotransferase domain.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
CC EMBL; BX936398; CAH23202.1; -; Genomic_DNA.
CC SMR; Q663R1; 1-239, 243-608.
CC GenomeReviews; BX936398 GR; YPTB3964.
CC HAMAP; MF_00164; -; 1.
CC InterPro; IPR000583; GATase_2.
CC InterPro; IPR005855; GlnS_trans.
CC InterPro; IPR001347; SIS.
CC Pfam; PF00310; GATase_2; 1.
CC TIGRFAMs; TIGR01135; glms; 1.
CC PROSITE; PS00443; GATASE TYPE II; 1.
CC AminoTransferase; Complete proteome; Glutamine amidotransferase;
CC Transferase.
FT INIT_MET 0 0 By similarity.
FT CHAIN 1 608 Glucosamine--fructose-6-phosphate
FT aminotransferase [isomerizing].
FT /FTID=PRO_0000135419
FT REGION 1 240 Glutamine amidotransferase.
FT ACT_SITE 1 1 GATase (By similarity).
FT ACT_SITE 603 603 Isomerization Fru-6P (By similarity).
SQ SEQUENCE 608 AA; 66394 MW; 1655FD0ADB16CCD6 CRC64;

Query Match
Best Local Similarity 86.7%; Score 2677; DB 1; Length 608;
Matches 517; Conservative 44; Mismatches 47; Indels 0; Gaps 0;

QY 2 CGIVGAIQAQRDVAREILLEGRLRLEGRYGYDSAGLAVVDAEGHMTLRRLGKVMQLAAAE 61
Db 1 CGIVGAVQAQRDIAEILLEGRLRLEGRYGYDSAGLAVVDSGHLTRLRRVGVKVALSDAAEK 60

QY 62 HPLHGGTGTIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEPLREELKARGTYFVS 121
Db 61 QDLHGGTGTIAHTRWATHGEPSEANAPHVSDYSIVVHNGIIEHNEPLRELLISRGYRFS 120

QY 122 ETDETEVIAHLVNWELKOGGTLRREAVRAIPQLRGAYGTVIMDSRHPDPTLLAARSGSPVLI 181
Db 121 ETDETEVIAHLVHVEQOQGSLLVVKRVIPLQLRGAYGTVVWDSRDLRLIARSGSPVLI 180

QY 182 GLGNGENFIASDQALLPVTRRRFIFLEEGDIAETRRSVNIFDKTGAEVKRODIESNLOY 241
Db 181 GCGVGENFIASDQALLPVTRRRFIFLEEGDVVEVTRRSISIFDKQNAIERPEIESQVQY 240

QY 242 DAGDKGIRYHYMQEIIYEOPNAIKNTLTGRISHGQVDLSGLGNADLLSKVEHIQILAC 301
Db 241 DAGDKGIRYHYMQEIIYEOPMAIKNTLEGLRSHGMIDLSELGPKADALLAEVQHIQIAC 300

QY 302 GTSYNSGMWSRYWPFESLAGIPCDVEIASPEFRYKSAVRNSLMITLSQSGETADTLAAGR 361
Db 301 GTSYNSGMWSRYWPFESLAGVPCDVEIASPEFRYKSAVRPNLSLITLSQSGETADTLAALR 360

QY 362 LSKELGYLGSALICNVPGSSLVRESLALMTNAGTIGVASTKFTTQTLVLLMLVAKLS 421
Db 361 LSKELGYLGSALICNVAGSSLVRESLALMTKAGTIGVASTKFTTQTLVLLMLVGRIG 420
```

```
Qy 422 RLKGLDASIEHDI VHGLQALPSRIEQMLSODKRIEALAEDEFSDKHHALFLGRGDQYPIAL 481
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
421 KLGADASLEHDI VHGLQALPSRIEQMLSODKRIEALAEDEFSDKHHALFLGRGDQYPIAM 480
Qy 482 EGALKLKEISYIIHAEAYAAGELKHGPIALIDADMPVIVVAPNNELLEKLKSNIEEVRARG 541
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
481 EGALKLKEISYIIHAEAYAAGELKHGPIALIDADMPVIVVAPNNELLEKLKSNIEEVRARG 540
Qy 542 GOLYVFADQDAGFVSSDNMHI IEMPHVEEVIAP IFTVPLQLLAYHVALIKGTVDQPRN 601
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
541 GLYVFADQDAGFTDSEGMKII QPHVEEIIAP IFTVPLQLLSYHVALIKGTVDQPRN 600
Qy 602 LAKSVTVE 609
Db :|||||:
601 LAKSVTVE 608
```

Search completed: August 7, 2006, 09:24:37
Job time : 157 secs

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 7, 2006, 09:25:52 ; Search time 24 Seconds
(without alignments)
2221.090 Million cell updates/sec

Title: US-10-612-779-2

Perfect score: 3086

Sequence: 1 MCGIVGAIQRDVAEILLEG.....LIRGTDVOPRLAKSVTVE 609

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Databases : Issued Patents AA:*

- 1: /EMC Cellerai_SIDS3/ptodata/2/iaa/5_COMB.pdp.*
- 2: /EMC Cellerai_SIDS3/ptodata/2/iaa/6_COMB.pdp.*
- 3: /EMC Cellerai_SIDS3/ptodata/2/iaa/7_COMB.pdp.*
- 4: /EMC Cellerai_SIDS3/ptodata/2/iaa/H_COMB.pdp.*
- 5: /EMC Cellerai_SIDS3/ptodata/2/iaa/FCUS_COMB.pdp.*
- 6: /EMC Cellerai_SIDS3/ptodata/2/iaa/RE_COMB.pdp.*
- 7: /EMC Cellerai_SIDS3/ptodata/2/iaa/backfiles1.pdp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|----------------------|
| 1 | 3086 | 100.0 | 609 | 2 | US-09-115-475-16 |
| 2 | 3080 | 99.8 | 609 | 2 | US-09-115-475-28 |
| 3 | 3080 | 99.8 | 609 | 2 | US-09-115-475-31 |
| 4 | 3079 | 99.8 | 609 | 2 | US-09-115-475-25 |
| 5 | 3071 | 99.5 | 609 | 2 | US-09-115-475-19 |
| 6 | 3068 | 99.4 | 609 | 2 | US-09-115-475-22 |
| 7 | 2876 | 93.2 | 651 | 2 | US-09-489-039A-10691 |
| 8 | 2510.5 | 81.4 | 622 | 2 | US-09-543-681A-7522 |
| 9 | 1995 | 64.6 | 616 | 2 | US-09-252-991A-19164 |
| 10 | 1780.5 | 57.7 | 613 | 2 | US-09-328-352-5065 |
| 11 | 1492.5 | 48.4 | 612 | 2 | US-09-540-236-2182 |
| 12 | 1362.5 | 44.2 | 572 | 2 | US-09-902-540-16012 |
| 13 | 1357.5 | 44.0 | 592 | 1 | US-08-599-171A-30 |
| 14 | 1357.5 | 44.0 | 592 | 1 | US-08-646-590B-30 |
| 15 | 1357.5 | 44.0 | 592 | 2 | US-09-069-226-30 |
| 16 | 1357.5 | 44.0 | 592 | 2 | US-09-412-184-30 |
| 17 | 1155 | 37.4 | 611 | 2 | US-09-438-185A-970 |
| 18 | 1137 | 36.8 | 655 | 2 | US-09-107-532A-4425 |
| 19 | 1117.5 | 36.2 | 602 | 2 | US-09-583-110-4085 |
| 20 | 1117.5 | 36.2 | 635 | 2 | US-09-107-433-5022 |
| 21 | 1115 | 36.2 | 619 | 2 | US-09-134-001C-4248 |
| 22 | 1103.5 | 35.8 | 682 | 1 | US-08-911-445-3 |
| 23 | 1103.5 | 35.8 | 682 | 2 | US-09-182-983-3 |
| 24 | 1103.5 | 35.8 | 682 | 2 | US-09-771-838A-3 |
| 25 | 1103.5 | 35.8 | 682 | 2 | US-09-949-016-6591 |
| 26 | 1089.5 | 35.3 | 680 | 2 | US-09-731-166-6 |

| | | | | | | |
|----|-------|------|-----|---|----------------------|-------------------|
| 27 | 960.5 | 31.1 | 615 | 1 | US-08-911-445-2 | Sequence 2, Appli |
| 28 | 960.5 | 31.1 | 615 | 2 | US-09-182-983-2 | Sequence 2, Appli |
| 29 | 960.5 | 31.1 | 615 | 2 | US-09-771-838A-2 | Sequence 2, Appli |
| 30 | 846.5 | 27.4 | 640 | 2 | US-09-134-000C-3871 | Sequence 3871, Ap |
| 31 | 745 | 24.1 | 487 | 2 | US-09-248-796A-17250 | Sequence 17250, A |
| 32 | 704 | 22.8 | 365 | 2 | US-09-949-016-11214 | Sequence 11214, A |
| 33 | 677 | 21.9 | 405 | 2 | US-09-710-279-232 | Sequence 232, App |
| 34 | 612.5 | 19.8 | 422 | 2 | US-09-270-767-45187 | Sequence 45187, A |
| 35 | 574 | 18.6 | 486 | 1 | US-08-911-445-18 | Sequence 18, Appl |
| 36 | 574 | 18.6 | 486 | 2 | US-09-182-983-18 | Sequence 18, Appl |
| 37 | 574 | 18.6 | 486 | 2 | US-09-771-838A-18 | Sequence 18, Appl |
| 38 | 573 | 18.6 | 425 | 1 | US-08-911-445-1 | Sequence 1, Appli |
| 39 | 573 | 18.6 | 425 | 2 | US-09-182-983-1 | Sequence 1, Appli |
| 40 | 573 | 18.6 | 425 | 2 | US-09-771-838A-1 | Sequence 320, App |
| 41 | 555 | 18.0 | 338 | 2 | US-09-710-279-320 | Sequence 1041, Ap |
| 42 | 496 | 16.1 | 307 | 2 | US-09-198-452A-1041 | Sequence 1040, Ap |
| 43 | 360.5 | 11.7 | 161 | 2 | US-09-198-452A-1040 | Sequence 18074, A |
| 44 | 332 | 10.8 | 708 | 2 | US-09-252-991A-18074 | Sequence 18074, A |
| 45 | 320 | 10.4 | 182 | 2 | US-09-198-452A-1042 | Sequence 1042, Ap |

ALIGNMENTS

RESULT 1

US-09-115-475-16
; Sequence 16, Application US/09115475
; Patent No. 6372457

; GENERAL INFORMATION:

; APPLICANT: Berry, Alan

; APPLICANT: Burlingame, Richard P.

; APPLICANT: Millis, James R.

; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE

; FILE REFERENCE: 3161-18-C1

; CURRENT APPLICATION NUMBER: US/09/115,475

; CURRENT FILING DATE: 1998-07-15

; EARLIER APPLICATION NUMBER: PCT/US98/00800

; EARLIER FILING DATE: 1998-01-14

; EARLIER APPLICATION NUMBER: 60/035,494

; EARLIER FILING DATE: 1997-01-14

; NUMBER OF SEQ ID NOS: 31

; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 16

; LENGTH: 609

; TYPE: PRT

; ORGANISM: Escherichia coli

US-09-115-475-16

Query Match 100.0%; Score 3086; DB 2; Length 609;
Best Local Similarity 100.0%; Pred. No. 3.7e-295;
Matches 609; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|-----|---|-------------------|
| Qy | 1 | MCGIVGAIQRDVAEILLEGRLRLEVRGYDSAGLVVDAEGHMTLRLRLGKQVQLAAAE 60 | Sequence 2, Appli |
| Db | 1 | MCGIVGAIQRDVAEILLEGRLRLEVRGYDSAGLVVDAEGHMTLRLRLGKQVQLAAAE 60 | Sequence 2, Appli |
| Qy | 61 | EHPHGGTGTAHTRWATHGEPSEVNAHPVSEHVVVHNGIIEHPELRELKARGYTFV 120 | Sequence 18, Appl |
| Db | 61 | EHPHGGTGTAHTRWATHGEPSEVNAHPVSEHVVVHNGIIEHPELRELKARGYTFV 120 | Sequence 18, Appl |
| Qy | 121 | SETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPV 180 | Sequence 11214, A |
| Db | 121 | SETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPV 180 | Sequence 11214, A |
| Qy | 181 | IGLGMGNFIASQLALLPVTRRFIFLEGGDIAEITRRSVNIIDKTAQVVRQDIESNLQ 240 | Sequence 232, App |
| Db | 181 | IGLGMGNFIASQLALLPVTRRFIFLEGGDIAEITRRSVNIIDKTAQVVRQDIESNLQ 240 | Sequence 232, App |
| Qy | 241 | YDAGDKGIYRHYWQKEIYEQNPAINKTLTGRIHSGQVDLSLGNADLLSKVEHIQILA 300 | Sequence 45187, A |
| Db | 241 | YDAGDKGIYRHYWQKEIYEQNPAINKTLTGRIHSGQVDLSLGNADLLSKVEHIQILA 300 | Sequence 45187, A |
| Qy | 301 | CGTSYNSGMVSRVWFESLAGIPCDVEIASEFRYKRSVRRNSLMITLSQSGETADTLA 360 | Sequence 1042, Ap |

Db 301 CGTSYNSGMVSRWYFESLAGIPCDVEIASBFYRKSAVRNSLMTLSQSGETADTLA 360
Qy 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQTLTVLLMLVAKL 420
Db 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQTLTVLLMLVAKL 420
Qy 421 SRLKGLDASIEHDIYHGLQALPSRIEOMLSQDKRIEALAEFSDKHHALFLGRGDOYPTA 480
Db 421 SRLKGLDASIEHDIYHGLQALPSRIEOMLSQDKRIEALAEFSDKHHALFLGRGDOYPTA 480
Qy 481 LEGALKLKEISYIHAEEYAAGELKHGPLALIDADMPVIVVAPNNLELKLKSNIEEVRAR 540
Db 481 LEGALKLKEISYIHAEEYAAGELKHGPLALIDADMPVIVVAPNNLELKLKSNIEEVRAR 540
Qy 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPIFTYVPLQLLAYHVALIKGTDVDDQPR 600
Db 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPIFTYVPLQLLAYHVALIKGTDVDDQPR 600
Qy 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 2

US-09-115-475-28
; Sequence 28, Application US/09115475
; Patent No. 6372457
; GENERAL INFORMATION:
; APPLICANT: Berry, Alan
; APPLICANT: Burlingame, Richard P.
; APPLICANT: Millis, James R.
; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE
; FILE REFERENCE: 3161-18-C1
; CURRENT APPLICATION NUMBER: US/09/115,475
; CURRENT FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: PCT/US98/00800
; EARLIER FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/035,494
; EARLIER FILING DATE: 1997-01-14
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-115-475-28

Query Match 99.8%; Score 3080; DB 2; Length 609;
Best Local Similarity 99.8%; Pred. No. 1.5e-294;
Matches 608; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MCGIVGAIAQRDVAEILLEGRLRLEYRGYDSAGLAVVDAEGHMTLRLRLKGVMQLAAAE 60
Db 1 MCGIVGAIAQRDVAEILLEGRLRLEYRGYDSAGLAVVDAEGHMTLRLRLKGVMQLAAAE 60
Qy 61 EPHLHGCTGTAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHPELREELKARGYTFV 120
Db 61 EPHLHGCTGTAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHPELREELKARGYTFV 120
Qy 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPLV 180
Db 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPLV 180
Qy 181 IGLGMGENFIASDQALLPVTRRFIFLEEGDIAETRRSVNIFDKTGAEVKRDIESNLQ 240
Db 181 IGLGMGENFIASDQALLPVTRRFIFLEEGDIAETRRSVNIFDKTGAEVKRDIESNLQ 240
Qy 241 YDAGDKGIYRHYMKOEIYEOPNAIKNTLTGRISHGOVDLSELGPNADLLSKVHEHIQILA 300
Db 241 YDAGDKGIYRHYMKOEIYEOPNAIKNTLTGRISHGOVDLSELGPNADLLSKVHEHIQILA 300
Qy 301 CGTSYNSGMVSRWYFESLAGIPCDVEIASBFYRKSAVRNSLMTLSQSGETADTLA 360

Db 301 CGTSYNSGMVSRWYFESLAGIPCDVEIASBFYRKSAVRNSLMTLSQSGETADTLA 360
Qy 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQTLTVLLMLVAKL 420
Db 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQTLTVLLMLVAKL 420
Qy 421 SRLKGLDASIEHDIYHGLQALPSRIEOMLSQDKRIEALAEFSDKHHALFLGRGDOYPTA 480
Db 421 SRLKGLDASIEHDIYHGLQALPSRIEOMLSQDKRIEALAEFSDKHHALFLGRGDOYPTA 480
Qy 481 LEGALKLKEISYIHAEEYAAGELKHGPLALIDADMPVIVVAPNNLELKLKSNIEEVRAR 540
Db 481 LEGALKLKEISYIHAEEYAAGELKHGPLALIDADMPVIVVAPNNLELKLKSNIEEVRAR 540
Qy 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPIFTYVPLQLLAYHVALIKGTDVDDQPR 600
Db 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPIFTYVPLQLLAYHVALIKGTDVDDQPR 600
Qy 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 3

US-09-115-475-31
; Sequence 31, Application US/09115475
; Patent No. 6372457
; GENERAL INFORMATION:
; APPLICANT: Berry, Alan
; APPLICANT: Burlingame, Richard P.
; APPLICANT: Millis, James R.
; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE
; FILE REFERENCE: 3161-18-C1
; CURRENT APPLICATION NUMBER: US/09/115,475
; CURRENT FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: PCT/US98/00800
; EARLIER FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/035,494
; EARLIER FILING DATE: 1997-01-14
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-115-475-31

Query Match 99.8%; Score 3080; DB 2; Length 609;
Best Local Similarity 99.8%; Pred. No. 1.5e-294;
Matches 608; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MCGIVGAIAQRDVAEILLEGRLRLEYRGYDSAGLAVVDAEGHMTLRLRLKGVMQLAAAE 60
Db 1 MCGIVGAIAQRDVAEILLEGRLRLEYRGYDSAGLAVVDAEGHMTLRLRLKGVMQLAAAE 60
Qy 61 EPHLHGCTGTAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHPELREELKARGYTFV 120
Db 61 EPHLHGCTGTAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHPELREELKARGYTFV 120
Qy 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPLV 180
Db 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPLV 180
Qy 181 IGLGMGENFIASDQALLPVTRRFIFLEEGDIAETRRSVNIFDKTGAEVKRDIESNLQ 240
Db 181 IGLGMGENFIASDQALLPVTRRFIFLEEGDIAETRRSVNIFDKTGAEVKRDIESNLQ 240
Qy 241 YDAGDKGIYRHYMKOEIYEOPNAIKNTLTGRISHGOVDLSELGPNADLLSKVHEHIQILA 300
Db 241 YDAGDKGIYRHYMKOEIYEOPNAIKNTLTGRISHGOVDLSELGPNADLLSKVHEHIQILA 300
Qy 301 CGTSYNSGMVSRWYFESLAGIPCDVEIASBFYRKSAVRNSLMTLSQSGETADTLA 360

```
Db 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASEFRYKSAVRNSLMIITLSQGETADTLAGL 360
Qy 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFPTTQLTVLMLVAKL 420
Db 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFPTTQLTVLMLVAKL 420
Qy 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEDFSCKHHALFLGRGDQYPIA 480
Db 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEDFSCKHHALFLSRGDQYPIA 480
Qy 481 LEGALKLKEISYTHAEAYAAAGELKHGPLALIDAMPVIVVAPNNELLEKLKSNIEEVRAR 540
Db 481 LEGALKLKEISYTHAEAYAAAGELKHGPLALIDAMPVIVVAPNNELLEKLKSNIEEVRAR 540
Qy 541 GGQLYVPADQAGFVSSDNHIIEMPHVEEVIAPIFYTVPLOLLAYHVALIKGTDVQDPR 600
Db 541 GGQLYVPADQAGFVSSDNHIIEMPHVEEVIAPIFYTVPLOLLAYHVALIKGTDVQDPR 600
Qy 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609
```

```
RESULT 4
US-09-115-475-25
; Sequence 25, Application US/09115475
; Patent No. 6372457
; GENERAL INFORMATION:
; APPLICANT: Berry, Alan
; APPLICANT: Burlingame, Richard P.
; APPLICANT: Millis, James R.
; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE
; FILE REFERENCE: 3161-18-C1
; CURRENT APPLICATION NUMBER: US/09/115,475
; CURRENT FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: PCT/US98/00800
; EARLIER FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/035,494
; EARLIER FILING DATE: 1997-01-14
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
; ORGANISM: Escherichia coli
US-09-115-475-25
```

```
Query Match 99.8%; Score 3079; DB 2; Length 609;
Best Local Similarity 99.8%; Pred. No. 1.8e-294;
Matches 608; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MCGIVGAIARDVAEILLEGRLRLEYRGYDSAGLAVVDAEGHMTLRLRLGKVQMLAQAAE 60
Db 1 MCGIVGAIARDVAEILLEGRLRLEYRGYDSAGLAVVDAEGHMTLRLRLGKVQMLAQAAE 60

Qy 61 EPHLHGCTGTAHTRWATHGEPSEVNAHPHVSEHIVVHNGIIEHHEPLREELKARGYTFV 120
Db 61 EPHLHGCTGTAHTRWATHGEPSEVNAHPHVSEHIVVHNGIIEHHEPLREELKARGYTFV 120

Qy 121 SETDTEVIAHLVNWELKQGGLTREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPLV 180
Db 121 SETDTEVIAHLVNWELKQGGLTREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPLV 180

Qy 181 IGLGMGFNIASDQALLPVTRRPIFLEEGDIAETRISVNIIDFKTGAEVKRDIESNLQ 240
Db 181 IGLGMGFNIASDQALLPVTRRPIFLEEGDIAETRISVNIIDFKTGAEVKRDIESNLQ 240

Qy 241 YDAGDKGIYRHYMKEIYEOPNAIKNTLTGRISHGQVDLSLGNADLLSKVEHIQILA 300
Db 241 YDAGDKGIYRHYMKEIYEOPNAIKNTLTGRISHGQVDLSLGNADLLSKVEHIQILA 300

Qy 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASEFRYKSAVRNSLMIITLSQGETADTLAGL 360
```

```
Db 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASEFRYKSAVRNSLMIITLSQGETADTLAGL 360
Qy 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFPTTQLTVLMLVAKL 420
Db 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFPTTQLTVLMLVAKL 420
Qy 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEDFSCKHHALFLGRGDQYPIA 480
Db 421 SRLKGLDASIEHDI VHGLQALPSRIEQMLSDQKRIEALAEDFSCKHHALFLSRGDQYPIA 480
Qy 481 LEGALKLKEISYTHAEAYAAAGELKHGPLALIDAMPVIVVAPNNELLEKLKSNIEEVRAR 540
Db 481 LEGALKLKEISYTHAEAYAAAGELKHGPLALIDAMPVIVVAPNNELLEKLKSNIEEVRAR 540
Qy 541 GGQLYVPADQAGFVSSDNHIIEMPHVEEVIAPIFYTVPLOLLAYHVALIKGTDVQDPR 600
Db 541 GGQLYVPADQAGFVSSDNHIIEMPHVEEVIAPIFYTVPLOLLAYHVALIKGTDVQDPR 600
Qy 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609
```

```
RESULT 5
US-09-115-475-19
; Sequence 19, Application US/09115475
; Patent No. 6372457
; GENERAL INFORMATION:
; APPLICANT: Berry, Alan
; APPLICANT: Burlingame, Richard P.
; APPLICANT: Millis, James R.
; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE
; FILE REFERENCE: 3161-18-C1
; CURRENT APPLICATION NUMBER: US/09/115,475
; CURRENT FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: PCT/US98/00800
; EARLIER FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/035,494
; EARLIER FILING DATE: 1997-01-14
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
; ORGANISM: Escherichia coli
US-09-115-475-19
```

```
Query Match 99.5%; Score 3071; DB 2; Length 609;
Best Local Similarity 99.5%; Pred. No. 1.1e-293;
Matches 606; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MCGIVGAIARDVAEILLEGRLRLEYRGYDSAGLAVVDAEGHMTLRLRLGKVQMLAQAAE 60
Db 1 MCGIVGAIARDVAEILLEGRLRLEYRGYDSAGLAVVDAEGHMTLRLRLGKVQMLAQAAE 60

Qy 61 EPHLHGCTGTAHTRWATHGEPSEVNAHPHVSEHIVVHNGIIEHHEPLREELKARGYTFV 120
Db 61 EPHLHGCTGTAHTRWATHGEPSEVNAHPHVSEHIVVHNGIIEHHEPLREELKARGYTFV 120

Qy 121 SETDTEVIAHLVNWELKQGGLTREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPLV 180
Db 121 SETDTEVIAHLVNWELKQGGLTREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPLV 180

Qy 181 IGLGMGFNIASDQALLPVTRRPIFLEEGDIAETRISVNIIDFKTGAEVKRDIESNLQ 240
Db 181 IGLGMGFNIASDQALLPVTRRPIFLEEGDIAETRISVNIIDFKTGAEVKRDIESNLQ 240

Qy 241 YDAGDKGIYRHYMKEIYEOPNAIKNTLTGRISHGQVDLSLGNADLLSKVEHIQILA 300
Db 241 YDAGDKGIYRHYMKEIYEOPNAIKNTLTGRISHGQVDLSLGNADLLSKVEHIQILA 300

Qy 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASEFRYKSAVRNSLMIITLSQGETADTLAGL 360
```

Db 301 CGTSYNGMSRYWFESLAGIPCDVEIASBFYRKSAVRNSLMTLSQSGETADTLAGL 360
QY 361 RLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQTLTVLLMLVAKL 420
Db 361 RLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQTLTVLLMLVAKL 420
QY 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMLSDQKRIEALAEEDFSKHHALFLGRGDQYPIA 480
Db 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMLSDQKRIEALAEEDFSKHHALFLGRGDQYPIA 480
QY 481 LEGALKLKEISYIHAEAYAGELKHGPLALIDAMPVIVVAPNNELLEKLKSNIEVRAR 540
Db 481 LEGALKLKEISYIHAEAYAGELKHGPLALIDAMPVIVVAPNNELLEKLKSNIEVRAR 540
QY 541 GGQLYVFADQDAGFVSSDNNHIIEMPHVEEVIAPFYFVPLQLLAYHVALIKGTDVQDPR 600
Db 541 GGQLYVFADQDAGFVSSDNNHIIEMPHVEEVIAPFYFVPLQLLAYHVALIKGTDVQDPR 600
QY 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 6

US-09-115-475-22
; Sequence 22, Application US/09115475
; Patent No. 6372457
; GENERAL INFORMATION:
; APPLICANT: Berry, Alan
; APPLICANT: Burlingame, Richard P.
; APPLICANT: Mullis, James R.
; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE
; FILE REFERENCE: 3161-18-C1
; CURRENT APPLICATION NUMBER: US/09/115,475
; CURRENT FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: PCT/US98/00800
; EARLIER FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/035,494
; EARLIER FILING DATE: 1997-01-14
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-115-475-22

Query Match 99.4%; Score 3068; DB 2; Length 609;
Best Local Similarity 99.5%; Pred. No. 2.2e-293;
Matches 606; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 MCGIVGAIARDVAEILLEGRLRLEYRGYDSAGLAVVDAEGHMTLRRLRGKQVQLAAQAE 60
Db 1 MCGIVGAIARDVAEILLEGRLRLEYRGYDSAGLAVVDAEGHMTLRRLRGKQVQLAAQAE 60
QY 61 EPHLGGTGTIAHTRWATHGEPSEVNAHPHVSEHIVVHNGIIEHNEPLREELKARGYTFV 120
Db 61 EPHLGGTGTIAHTRWATHGEPSEVNAHPHVSEHIVVHNGIIEHNEPLREELKARGYTFV 120
QY 121 SETDTEVIAHLVNWELKQGGLTREAVALRAIPQLRGAYGTIVMDSRHPDPTLLAARSGSPV 180
Db 121 SETDTEVIAHLVNWELKQGGLTREAVALRAIPQLRGAYGTIVMDSRHPDPTLLAARSGSPV 180
QY 181 IGLGMENFIASQDALLPVTTRRFIFLEEGDIAETITRRSVNIFDKTGAEVKQDIESNLQ 240
Db 181 IGLGMENFIASQDALLPVTTRRFIFLEEGDIAETITRRSVNIFDKTGAEVKQDIESNLQ 240
QY 241 YDAGDKGIYRHYMQKEIYEOPNAIKNTLTGRISHGQVDLSLGPNADELKSKVHIQILA 300
Db 241 YDAGDKGIYRHYMQKEIYEOPNAIKNTLTGRISHGQVDLSLGPNADELKSKVHIQILA 300
QY 301 CGTSYNGMSRYWFESLAGIPCDVEIASBFYRKSAVRNSLMTLSQSGETADTLAGL 360

Db 301 CGTSYNGMSRYWFESLAGIPCDVEIASBFYRKSAVRNSLMTLSQSGETADTLAGL 360
QY 361 RLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQTLTVLLMLVAKL 420
Db 361 RLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQTLTVLLMLVAKL 420
QY 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMLSDQKRIEALAEEDFSKHHALFLGRGDQYPIA 480
Db 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMLSDQKRIEALAEEDFSKHHALFLGRGDQYPIA 480
QY 481 LEGALKLKEISYIHAEAYAGELKHGPLALIDAMPVIVVAPNNELLEKLKSNIEVRAR 540
Db 481 LEGALKLKEISYIHAEAYAGELKHGPLALIDAMPVIVVAPNNELLEKLKSNIEVRAR 540
QY 541 GGQLYVFADQDAGFVSSDNNHIIEMPHVEEVIAPFYFVPLQLLAYHVALIKGTDVQDPR 600
Db 541 GGQLYVFADQDAGFVSSDNNHIIEMPHVEEVIAPFYFVPLQLLAYHVALIKGTDVQDPR 600
QY 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 7

US-09-489-039A-10691
; Sequence 10691, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709-2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10691
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10691

Query Match 93.2%; Score 2876; DB 2; Length 651;
Best Local Similarity 92.1%; Pred. No. 2.2e-274;
Matches 561; Conservative 29; Mismatches 19; Indels 0; Gaps 0;

QY 1 MCGIVGAIARDVAEILLEGRLRLEYRGYDSAGLAVVDAEGHMTLRRLRGKQVQLAAQAE 60
Db 43 MCGIVGAIARDVAEILLEGRLRLEYRGYDSAGLAVVDAEGHMTLRRLRGKQVQLAAQAE 102
QY 61 EPHLGGTGTIAHTRWATHGEPSEVNAHPHVSEHIVVHNGIIEHNEPLREELKARGYTFV 120
Db 103 EPHLGGTGTIAHTRWATHGEPSEVNAHPHVSEHIVVHNGIIEHNEPLRALLQSRGYVVF 162
QY 121 SETDTEVIAHLVNWELKQGGLTREAVALRAIPQLRGAYGTIVMDSRHPDPTLLAARSGSPV 180
Db 163 SETDTEVIAHLVNWELKQGGLTREAVALRAIPQLRGAYGTIVMDSRHPDPTLLAARSGSPV 222
QY 181 IGLGMENFIASQDALLPVTTRRFIFLEEGDIAETITRRSVNIFDKTGAEVKQDIESNLQ 240
Db 223 IGLGMENFIASQDALLPVTTRRFIFLEEGDIAETITRRSVVIFDKSAQVFKQIESNLQ 282
QY 241 YDAGDKGIYRHYMQKEIYEOPNAIKNTLTGRISHGQVDLSLGPNADELKSKVHIQILA 300
Db 283 YDAGDKGIYRHYMQKEIYEOPNAIKNTLTGRISHGEVDLSLGPNAEMLAQVEHIQIVA 342
QY 301 CGTSYNGMSRYWFESLAGIPCDVEIASBFYRKSAVRNSLMTLSQSGETADTLAGL 360
Db 343 CGTSYNGMSRYWFESLAGIPCDVEIASBFYRKSAVRNSLMTLSQSGETADTLAGL 402
QY 361 RLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQTLTVLLMLVAKL 420

```
Db 403 RLSKELGYLSLAICNVPGSSLVRESLALMTKAGTEIGVASTKFTTQLTVLLMLVAKL 462
Qy 421 SRLKGLDASTEHDI VHGLQALPSRIEOMLSQDKRIEALAEFSDKKHALFLGRGDQYPIA 480
Db 463 ARUKGQDASTEHDI VHGLQALPNRIEOMLSQDKRIEALAEFSDKKHALFLGRGDQYPIA 522
Qy 481 LEGALKLKEISYTHAEAYAGELKHGFLALIDAMPVIVVAPNNLELLEKLSNIEEVRAR 540
Db 523 MEGALKLKEISYTHAEAYAGELKHGFLALIDAMPVIVVAPNNLELLEKLSNIEEVRAR 582
Qy 541 GGQLYVPADQADGAFVSSDNMHIIMPHVEEVIAPIFTVPQLLQALYHVALIKGTDVDPQR 600
Db 583 GGELYVPADGAGFNGSDNMHIIMPHVEETIAPIFTVPQLLQALYHVALIKGTDVDPQR 642
Qy 601 NLAKSVTVE 609
Db 643 NLAKSVTVE 651

RESULT 8
US-09-543-681A-7522
; Sequence 7522, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543.681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7522
; LENGTH: 622
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-7522

Query Match 81.4%; Score 2510.5; DB 2; Length 622;
Best Local Similarity 81.1%; Pred. No. 2.4e-238;
Matches 494; Conservative 46; Mismatches 68; Indels 1; Gaps 1;

Qy 1 MCGIVGAIAORDVAEILLEGRLRELYRGYSAGLAVVDAEGHMTLRLRLGKVMQLAQAE 60
Db 15 MCGIVGAIAORDVAEILLEGRLRELYRGYSAGLAVVDAEGHMTLRLRLGKVMQLAQAE 74
Qy 61 EPHLHGCTGTAHTRWATHGEPSEVNAHPHVSEHIVVHNGIIEHNEPLRBEELKARGYTFV 120
Db 75 KTVIGGTAHTRWATHGEPSEVNAHPHVSEHIVVHNGIIEHNEPLRBEELKARGYTFV 134
Qy 121 SETDTEVIAHLVNWELKOGGTLREAVLRAIPQLRGAYGTVMDSRHPDPTLLAARSGSPLV 180
Db 135 SQTDTTEVIAHLVNWELKOGGTLREAVLRAIPQLRGAYGTVMDSRHPDPTLLAARSGSPLV 194
Qy 181 IGLMGENFIASDQALLPVTRRRFIFLEEGDIAETRRSVNIPTKTAEVKRODIESNLQ 240
Db 195 VGLGVGENFLASDQALLPVTRRRFIFLEEGDIAETRRSVNIPTKTAEVKRODIESNLQ 254
Qy 241 YDAGDKGIYRHYMKEIYEOPNAIKNTLTGRISHGQVDLSLGNADDELLSKVEHIQILA 300
Db 255 YDAGDKGIYRHYMKEIYEOPNAIKNTLTGRISHGQVDLSLGNADDELLSKVEHIQILA 314
Qy 301 CGTSYNGMVSRYWFESLAGIPCDVEIASFRRYKSAVRNLSMITLSQSGETADTLAAGL 360
Db 315 CGTSYNGMVSRYWFESLAGIPCDVEIASFRRYKSAVRNLSMITLSQSGETADTLAAGL 374
Qy 361 RLSKELGYLSLAICNVPGSSLVRESLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 420
Db 375 RLSKELGYLSLAICNVAGSSLVRESFVLMTKAGAEIGVASTKFTTQLAVLLMVAAYM 434
Qy 421 SRLKGLDASTEHDI VHGLQALPSRIEOMLSQDKRIEALAEFSDKKHALFLGRGDQYPIA 480
```

```
Db 435 GRITKGVEA-LEQQVAMALHALPSRIEOMLSQDKRIEALAEFSEKSHALFLGRGDQYPIA 493
Qy 481 LEGALKLKEISYTHAEAYAGELKHGFLALIDAMPVIVVAPNNLELLEKLSNIEEVRAR 540
Db 494 VEGALKLKEISYTHAEAYAGELKHGFLALIDAMPVIVVAPNNLELLEKLSNIEEVRAR 553
Qy 541 GGQLYVPADQADGAFVSSDNMHIIMPHVEEVIAPIFTVPQLLQALYHVALIKGTDVDPQR 600
Db 554 GGLLYVPADQADGAFSEETMKLISLPHVEELIAPIFTVPQLLQALYHVALIKGTDVDPQR 613
Qy 601 NLAKSVTVE 609
Db 614 NLAKSVTVE 622

RESULT 9
US-09-252-991A-19164
; Sequence 19164, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19164
; LENGTH: 616
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19164

Query Match 64.6%; Score 1995; DB 2; Length 616;
Best Local Similarity 62.7%; Pred. No. 1.7e-187;
Matches 383; Conservative 101; Mismatches 125; Indels 2; Gaps 2;

Qy 1 MCGIVGAIAORDVAEILLEGRLRELYRGYSAGLAVVDAEGHMTLRLRLGKVMQLAQAE 60
Db 6 MCGIVGAIAERNITPILIEGLKRELYRGYSAGVAVFDNEGRLQRCRRGVKVASLEEGLA 65
Qy 61 EPHLHGCTGTAHTRWATHGEPSEVNAHPHVSEHIVVHNGIIEHNEPLRBEELKARGYTF 119
Db 66 GTPLLGRLGTAHTRWATHGAPTEGNAHPHFSSDELAVVHNGIIEHNEPLRBEELKARGYTF 125
Qy 120 VSETDTEVIAHLVNWELKOGGTLREAVLRAIPQLRGAYGTVMDSRHPDPTLLAARSGSPL 179
Db 126 TSQDTEVIAHLVHLLHKKQSIGDUTLAKDAVKELHGAIGLAVISAAPDRIVAAARSGSPL 185
Qy 180 VIGLGMGENFIASDQALLPVTRRRFIFLEEGDIAETRRSVNIPTKTAEVKRODIESNL 239
Db 186 VIGLGMGENFIASDQALLPVTRRRFIFLEEGDIAETRRSVNIPTKTAEVKRODIESNL 245
Qy 240 QYDAGDKGIYRHYMKEIYEOPNAIKNTLTGRISHGQVDLSLGNADDELLSKVEHIQIL 299
Db 246 GABEADKGEYRHFMLKEIHEQPSVQVORTLEGRIGONQVMVESFGPQAAELFAKVRNVQIV 305
Qy 300 ACCTSNGMVSRYWFESLAGIPCDVEIASFRRYKSAVRNLSMITLSQSGETADTLAAGL 359
Db 306 ACCTSNGMVSRYWFESLAGIPCDVEIASFRRYKSAVRNLSMITLSQSGETADTLAAGL 365
Qy 360 LRLSKELGYLSLAICNVPGSSLVRESLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 419
Db 366 LRLSKELGYLSLAICNVATSSLVRESLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 425
Qy 420 LSKL-KGLDASIEHDI VHGLQALPSRIEOMLSQDKRIEALAEFSDKKHALFLGRGDQY 478
Db 426 IGQVQKRLADGVEAEVLDELRLPTRLGALAMNRTVEKVSSELPAEKHHTLFLGRGAQFP 485
```

| | | | |
|---|-----|---|-----|
| QY | 479 | I A L E G A L K L K E I S Y I H A E Y A A G E L K H G P L A L I D A D M P V I V V A P N N E L L E K L S N I E E V R | 538 |
| D b | 486 | V A L E G A L K L K E I S Y I H A E Y A P A G E L K H G P L A L V D S M P V V I V A P N N E L L E K L S N I Q E V R | 545 |
| QY | 539 | A R G G Q L Y V F A D Q D A G F V S S D N M H I I E M P H V E E V I A F I F Y T V P L Q L L A Y H V A L I K G T D V D Q | 598 |
| D b | 546 | A R G G E L V V F A D E G A G I E A G E T H V G M P H I G D V L S P I L Y T I P L Q L L S Y H V A V L K G T D V D Q | 605 |
| QY | 599 | P R N L A K S V T V E | 609 |
| D b | 606 | P R N L A K S V T V E | 616 |
| RESULT 10 | | | |
| US-09-328-352-5066 | | | |
| ; Sequence 5066, Application US/09328352 | | | |
| ; Patent No. 6562958 | | | |
| ; GENERAL INFORMATION: | | | |
| ; APPLICANT: Gary L. Breton et al. | | | |
| ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER | | | |
| ; FILE REFERENCE: GTC99-03PA | | | |
| ; CURRENT APPLICATION NUMBER: US/09/328,352 | | | |
| ; CURRENT FILING DATE: 1999-06-04 | | | |
| ; NUMBER OF SEQ ID NOS: 8252 | | | |
| ; SEQ ID NO 5066 | | | |
| ; LENGTH: 613 | | | |
| ; TYPE: PRT | | | |
| ; ORGANISM: Acinetobacter baumannii | | | |
| US-09-328-352-5066 | | | |
| Query Match 57.7%; Score 1780.5; DB 2; Length 613; | | | |
| Best Local Similarity 57.4%; Pred. No. 2.5e-166; | | | |
| Matches 353; Conservative 107; Mismatches 146; Indels 9; Gaps 5; | | | |
| QY | 1 | M C G I V G A I A Q R D V A E I L L E G L R R L E Y R G Y D S A G L A V V D A E G H M T R R L R L K G V Q M L A Q A A E | 60 |
| D b | 2 | M C G I V G V A E R C V T E I L E G L K E L Y R G Y D S A G V A L L N K Q - Q I L R E R R V G K V I N L E D A V A | 60 |
| QY | 61 | E H P L H G T G T I A H T R W A T H G E P S E V N A H P H V S E H I V V H N G I I E N H P L R E E L K A R G Y T V | 120 |
| D b | 61 | D Q O L T G T I G I A H T R W A T H G K T E N N A H P H M S G K V A V V H N G I I E N Y Q E L X D D L Q A L G Y V E T | 120 |
| QY | 121 | S E T D T E V I A H L V N W E L K O G G T L R E A V L R A I P O L R G A Y G T V I M D S R H P D T L L A A R S G S P L A | 180 |
| D b | 121 | S Q T D T E V A H L V A E A L K N D T S L L E A V E S V V P Q L K A Y A L G I I H S D Y P D E L I T V R E G S P L V | 180 |
| QY | 181 | I G L G M E N F T A S D Q L A L L P V T R R F I F L E E G D I A E I T R R S V N I P D K T G A E V K R Q D I E S N I Q | 240 |
| D b | 181 | I G V G I G E N F I S S D Q L A L L P V T N R F I Y L E E G D I A R L T R T S I E V F V K - G E R V E R P V K E L D A T | 239 |
| QY | 241 | Y D A G D K G I Y R H Y M O K E I Y E Q P N A I K N T L T G R I S H G Q V D L S E L G P N A D E L L S K V E H I Q I L A | 300 |
| D b | 240 | V S S A S K G E Y K H Y M L K E I Y E Q P E A I K O T I S Q A L D - G N N L R D D F L K D A E A D F S K L Q S Q I T A | 298 |
| QY | 301 | C G T S Y N S G M V S R Y W F E S L A G I P C D V E I A S E F R Y K S A V R N S I M I T L S O S G E T A D T L A G L | 360 |
| D b | 299 | C G T S Y H S G M I A K Y F O L G V P C Q V E I A S E F R Y S P V I E N T I Y I C I S O S G E T A D T L A A L | 358 |
| QY | 361 | R L S K E L G Y - - - - - L G S I A I C N V P G S S L V R E S D I A L M T N A G T E I G V A S T K A F T T Q L T V L I M | 415 |
| D b | 359 | R E T Q R A K A N N I D I Q T L I C N V A T S S W R E T D H L L T L A G P E I G V A S T K A F T T Q L A A L M L | 418 |
| QY | 416 | L V A K L S R L K G L D A S I E - H D I V H G L Q A L P S R I E Q M L S O D K R I E A L A B D F S D K H A L F L G R G | 474 |
| D b | 419 | L I L K I G Q V K Q R I S N V M I E A L E A R L M H S P K V I L D T L K Q D A E I L R S L S E L F V E K Q H C I F L G R G | 478 |
| QY | 475 | D Q Y P I A L E G A L K L K E I S Y I H A E Y A A G E L K H G P L A L I D A D M P V I V V A P N N E L L E K L S N I | 534 |
| D b | 479 | T H Y P I A L E G A L K L K E I S Y I H A E Y A A G E L K H G P L A L V D N E M P I V I L A P N D E M D L K L S N M | 538 |
| QY | 535 | E E V R A R G G Q L Y V F A D Q D A G F V S S D N M H I I E M P H V E E V I A F I F Y T V P L Q L L A Y H V A L I K G T | 594 |

| | | | |
|--|-----|---|-----|
| D b | 539 | E E V Q A R G G E L F V F A D E N S G V E K D R Q H V Q I P A V N E W L A P I I Y S V P V Q L L S Y H V A V L R C T | 598 |
| QY | 595 | D V D Q P R N L A K S V T V E | 609 |
| D b | 599 | D V D Q P R N L A K S V T V E | 613 |
| RESULT 11 | | | |
| US-09-540-236-2182 | | | |
| ; Sequence 2182, Application US/09540236 | | | |
| ; Patent No. 6673910 | | | |
| ; GENERAL INFORMATION: | | | |
| ; APPLICANT: Gary L. Breton et al. | | | |
| ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARI | | | |
| ; FILE REFERENCE: 2709.2005-001 | | | |
| ; CURRENT APPLICATION NUMBER: US/09/540,236 | | | |
| ; CURRENT FILING DATE: 2000-04-04 | | | |
| ; NUMBER OF SEQ ID NOS: 3840 | | | |
| ; SEQ ID NO 2182 | | | |
| ; LENGTH: 619 | | | |
| ; TYPE: PRT | | | |
| ; ORGANISM: M. catarrhalis | | | |
| US-09-540-236-2182 | | | |
| Query Match 48.4%; Score 1492.5; DB 2; Length 619; | | | |
| Best Local Similarity 50.3%; Pred. No. 6.8e-138; | | | |
| Matches 309; Conservative 112; Mismatches 186; Indels 7; Gaps 7; | | | |
| QY | 1 | M C G I V G A I - A Q R D V A E I L L E G L R R L E Y R G Y D S A G L A V V D A E G H M T R R L R L K G V Q M L A Q A A | 59 |
| D b | 8 | M C G I V G A I R T H N V V D F L T D G L K R L E Y R G Y D S S G I A - V Q T E S G I R R V R V R G R V A L M E E A A | 66 |
| QY | 60 | E H P L H G T G T I A H T R W A T H G E P S E V N A H P H V S E H I V - V H N G I I E N H P L R E E L K A R G Y T | 118 |
| D b | 67 | K Q K Q I H G L T G I C H T R W A T H G V T K P N A H P H I S A G L I S V V H N G I I E N F E T E R T R L Q H L D Y E | 126 |
| QY | 119 | F V S E T D T E V I A H L V N W E - L K O G G T L R E A V L R A I P O L R G A Y G T V I M D S R H P D T L L A A R S G S | 177 |
| D b | 127 | F E S Q T D T E V I A H S V H Q E Y T T N G G D L Y H A V Q T A C S R F H G A Y A I A V A N D A P N Q M V I A R M G C | 186 |
| QY | 178 | P L V I G L G M E N F I A S D Q L A L L P V T R R F I F L E E G D I A E I T R R S V - N I F D K T G A E V K R Q D I E | 236 |
| D b | 187 | P L L V A F A G D N E V F A S D V S A V I A F T R R V T Y L E D G D I A L L Q A D G I O K L L D K D G N Q V N R S I K T | 246 |
| QY | 237 | S N L Q V D A G D K G I Y R H Y M O K E I Y E Q P N A I K N T L T G R I S H G Q V D L S E L G P N A D E L L S K V E H I | 296 |
| D b | 247 | S E M S I A S L E L G P Y S H F M Q K E I N E Q P R A V S D T A E I F L E G G F V - A D N F G E A P E I F A K I Q S I | 305 |
| QY | 297 | Q I L A C G T S Y N S G M V S R Y W F E S L A G I P C D V E I A S E F R Y K S A V R R N S L M I T L S O S G E T A D T | 356 |
| D b | 306 | K I L A C G T S Y A A L T G K Y L E S I A K I R C D V E I A S E Y R D V I S D P D E L V I T I S O S G E T L D T | 365 |
| QY | 357 | L A G L R L S K E L G Y L G S L A I C N V P G S S L V R E S D I A L M T N A G T E I G V A S T K A F T T Q L T V L I M L | 416 |
| D b | 366 | M E A L K Y A N A Q H R Y S L S I C N V M E S A L P R T S T L A I F T R A G A E I G V A S T K A F T T Q L V V L F C L | 425 |
| QY | 417 | V A K L S R L K G - I D A S I E H D I V H G L Q A L P S R I E Q M L S O D K R I E A L A B D F S D K H A L F L G R G D | 475 |
| D b | 426 | A V T L G I L R H G V D E S Q A Q Y A E D L R L L P G S I O H A L N L E P Q L A S W A Q S F A N K F S A L F L G R G I | 485 |
| QY | 476 | C Y P T A L E G A L K L K E I S Y I H A E Y A A G E L K H G P L A L I D A D M P V I V V A P N N E L L E K L S N I E | 535 |
| D b | 486 | H Y P T A L E G A L K L K E I Y I H A E S Y P A G E L K H G P L A L V D N E M P V V I S P N D G L L D K R Y A N N Q | 545 |
| QY | 536 | E V R A R G G Q L Y V F A D Q D A G F V S S D N M H I I E M P H V E E V I A F I F Y T V P L Q L L A Y H V A L I K G T D | 595 |
| D b | 546 | E V S A R G G E L F V L S D L D S D Y T A S E G V H I R T P R H I T G L S P I V H T I P V Q L L A Y H V A L V K G T D | 605 |
| QY | 596 | V D Q P R N L A K S V T V E | 609 |
| D b | 606 | V D K P R N L A K S V T V E | 619 |

RESULT 12
US-09-902-540-16012
; Sequence 16012, Application US/0902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 16012
; LENGTH: 572
; TYPE: PR1
; ORGANISM: Myxococcus xanthus
US-09-902-540-16012

Query Match 44.2%; Score 1362.5; DB 2; Length 572;
Best Local Similarity 50.0%; Pred. No. 4e-125;
Matches 279; Conservative 105; Mismatches 161; Indels 13; Gaps 7;
QY 59 ABEHPLHGGTGIATRWATHGEPSEVNAHPHSEHIVVHNGIENHEPRLBELKARGYT 118
DB 21 ADQPP--GNIGIGHTRWATHGRPSDENAHPTTYKDVAVVHNGIENHLSLKEQLRSRGHV 78
QY 119 FVSETDTEVIAHLVNNELKGGTLREAVLRAIPQLAGAYGTVMDSRHPDPTLLAARSGSP 178
DB 79 FSSETSEVFVHILSELENGLELPDAVRAIAQVKGTVLALVATSDPSRIVCTKDASP 138
QY 179 LVILGGMENFIASDQALLPVTRRFIFLEEGDIAETRISVNI FDKTGAEVKRODIESN 238
DB 139 MVILGGQGNFLASDVPALLETHTDRDFVYMBEGDLAVITAAAVDFNKGQKVNRPTRRID 198
QY 239 LQYDAGDKGIYRHYMOKIEYEQNAIKNTLTGR--ISHGQVDLSELGPNADLLSKVEHI 296
DB 199 WTPMAEKGGKHFMHKEIWEQPRVADTLRGRMLLSEGDVHPEGNLSAEKVRS-LTKI 257
QY 297 QILACGTSYNSGMVSRWYFESLAGIPCDVEIASEFRYKSAVRNSLMTLSQSGETADT 356
DB 258 TILACGTSWHSGVAGKHMIESLARLPEVELASEFRYRDPIDVGTHLAIAISQSGETADT 317
QY 357 LAGLRSLKELGYLGSLAICNPVGSLSVRESDLALMTNAGTEIGVASTKFTTOLTLLML 416
DB 318 LAAPKEAKARG-ATANAICNVIGSAMTREADFSVLTNAGPEIGVASTKFTTQLVALFLL 376
QY 417 VAKLSRLKG---LDASIEHDIHVHGLQALPSRIEQLMSQDKRIEALAEFSDKHHLFLGR 473
DB 377 AVKLGMRGTLSPAAQEH--LTQLTKVPKMIEDVLKCEPAVTRVSREYMSQDFLFLGR 434
QY 474 GDQYPTALEGALKKEISYTHABAYAGELKHGCPALIDADMPIVIVAPN--NELLEKLEK 531
DB 435 GMPHPVALLGALKKEISYTHAEGYAGGEMKGPALIDEKMPVVIAPKQPHVAYEKII 494
QY 532 SNIEEVRAGGQLYVFADQAGVSSDNHMIEMPHVEEVIAPFVTPQLLAYHVALI 591
DB 495 GNIEEVRAGGKVIAlDEDEHVA TLADQVTRIPACALLAPVAVATIFLQLLAYHVALI 554
QY 592 KGTVDQPNRLAKSVTVE 609
DB 555 RGNVDQPNRLAKSVTVE 572

RESULT 13
US-08-599-171A-30
; Sequence 30, Application US/08599171A
; Patent No. 5814473

; GENERAL INFORMATION:
; APPLICANT: WARREN, Patrick V.
; TITLE OF INVENTION: TRANSAMINASES AND AMINOTRANSFERASES
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/599,171A
; FILING DATE: Concurrently
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: HERRON, CHARLES J.
; REGISTRATION NUMBER: 28,019
; REFERENCE/DOCKET NUMBER: 331400-38
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 592 AMINO ACIDS
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
US-08-599-171A-30

Query Match 44.0%; Score 1357.5; DB 1; Length 592;
Best Local Similarity 46.5%; Pred. No. 1.3e-124;
Matches 286; Conservative 113; Mismatches 187; Indels 29; Gaps 9;
QY 1 MCGIVCAIAQRDVA-EILLEGRLREYRCYDSAGLAVDAEGHMTLRRLRGKVMQLAQA 59
DB 1 MCGIVGVG-RDLALPVLIGALERLERYGYSAGVALIE-DGKLIVKKKGIKRELKVAL 58
QY 60 EEPHLHGGTGIATRWATHGEPSEVNAHPHVS--HIVVHNGIENHEPRLBELKARGY 117
DB 59 WKDVKAKTGIGHTRWATHGKPTDENAHPTDEKGEFAVVHNGIENYLELEKEELKEGV 118
QY 118 TVFSETDTEVIAHLVNNELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDPTLLAARSGS 177
DB 119 KFRSETDTEVIAHLIAKANYR--GDLLEAVLKTVKLGAFAVITVHEPNRLIGVKQGS 176
QY 178 PLVILGGMENFIASDQALLPVTRRFIFLEEGDIAETRISVNI FDKTGAEVKRODIES 237
DB 177 PLIVLGEGENFLASDIPAILPYTKIILVDDGEIADLTPTVTNVIYFEGEPVSKVEMIT 236
QY 238 NLQYDAGDKGIYRHYMOKIEYEQNAIKNTLTGRISHGQVDLSELGPNADL---LSKVE 294
DB 237 PNDLVSAEKGGKHPMLKEIYEQPKAINDTLKGFLS-----TEDAIFPKLKDPR 295
QY 295 HIQILACGTSYNSGMVSRWYFESLAGIPCDVEIASEFRYKSAVRNSLMTLSQSGETA 354
DB 286 RVLIIACGTSYHAGFVGKYWIERFAGVPTEVIYASEFRYADVPSDKDIVIGISQSGETA 345
QY 355 DTLAGLRSLKELGYLGSLAICNPVGSLSVRESDLALMTNAGTEIGVASTKFTTOLTLL 414
DB 346 DTKFALQSAKEKGF-TVGLVNVVSGAIDRESDFSLHTHAGPEIGVAATKFTTFAQTALY 404
QY 415 MLVAKLSRLKGLDASIEHDIHVHGLQALPSRIEQLMSQDKRIEALAEFSDKHHLFLGRG 474

Db 405 ALSVR-----ESEERENLRLEKVPESLVEQTINTAEVEKVAEKYMKKNWLYLGRY 457
QY 475 DOYPALLEGALKLKEISYIHAAYAGELKHGPIALIDAMPVIVVAPNNLEKLSNI 534
Db 458 LNYPIALEGALKLKEISYIHAEGYPAGEMKHGPIALIDENPVVVIAPKDRVYEKILSNV 517
QY 535 BEVRARGQLVVFADQDAGFVSSDNMHIEMPHVEEVIAPIFYVPLQLLAYHVALIKGT 594
Db 518 BEVLARKGRVISVGKGDTELKSKSESVMIBPKAEBIPFPLTVIPLQLFAYFIASKLGL 577
QY 595 DVDQPRNLAKSVTVE 609
Db 578 DVDQPRNLAKTVTVE 592
RESULT 14
US-08-646-590B-30
; Sequence 30, Application US/08646590B
; Patent No. 5962283
; GENERAL INFORMATION:
; APPLICANT: Warren, Patrick V.
; APPLICANT: Swanson, Ronald V.
; TITLE OF INVENTION: TRANSAMINASES AND AMINOTRANSFERASES
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,590B
; FILING DATE: 08-May-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/599,171
; FILING DATE: 09-FEB-1996
; PRIOR APPLICATION DATA: PCT/US97/01094
; APPLICATION NUMBER: PCT/US97/01094
; FILING DATE: 21-January-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Ph.D., Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 09010/017001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 592 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-646-590B-30
Query Match 44.0%; Score 1357.5; DB 1; Length 592;
Best Local Similarity 46.5%; Pred. No. 1.3e-124;
Matches 286; Conservative 113; Mismatches 187; Indels 29; Gaps 9;
QY 1 MCGIVGAIQADVA-BILLEGLRLRYGYDSAGLAVDAGHMTLRRLRGKVMQAA 59
Db 1 MCGIVGVG-RDALPVLIVGALERLEJRYGYDSAGVALIE-DKLIYEKKKGKIRELVAL 58
QY 60 BEHPLHGGTGIATRWATHGEPSEVNAHPHVSE--HIVVVHNGIIEHHEPLREELKARG 117

Db 59 WKQDYKAKTGIGHTRWATHGKETDENAHPTDEKGEFAVVHNGIIEYIELKEELKQGV 118
QY 118 TFVSETDETVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTIVIMDSRHPDPTLLAARSG 177
Db 119 KFRSETDETVIAHLIAKNYR--GDLEAVLTKVKLKGAFAPAVITVHPEPNRLIGVKQGS 176
QY 178 PLVIGLHGENFIASDQALLPVTRRFIFLEBGDTAEITRRSVNIFPDKTGAEVKKQDIES 237
Db 177 PLVIGLGEENFLASDIPAILFYTKKIIIVLDDGEIADLTPTDVTNINFEPEFVSEVMIT 236
QY 238 NLQYDAGDKGIYRHYMOKEIYBOPNAIKNTLTGRISHGQVDLSLSELGPNADEL--LSKVE 294
Db 237 PWDLSAEKGGFKHMLKEIYBOPKAINDTLKGFLS-----TEDAIPFKLKDRP 285
QY 295 HIQILACGTSYNSGMVSRYFESLAGIPCDVEIASEFRYKSAVRNLSMITLSQGETA 354
Db 286 RVLIIACGTSYHAGFVGKWIETRFAGVPTVIYASEPRYADVPSDKDVIIGISQGETA 345
QY 355 DTLAGLRLSKELGYLGLSLAICNVPOSSLVRESDLALMTNAGTEIGVASTKAPTQTOLTVLL 414
Db 346 DTKFALQSAKEKGAF-TVGLVNVVWSAIDRESDFSLSHTHAGPEIGVAATKTFTAQFTALY 404
QY 415 MELVAKLSRLKGLDASIEHDIHVHGLQALPSRIEOMLSQDKRIEALAEFSDKHHLFLGRG 474
Db 405 ALSVR-----ESEERENLRLEKVPESLVEQTINTAEVEKVAEKYMKKNWLYLGRY 457
QY 475 DOYPALLEGALKLKEISYIHAAYAGELKHGPIALIDAMPVIVVAPNNLEKLSNI 534
Db 458 LNYPIALEGALKLKEISYIHAEGYPAGEMKHGPIALIDENPVVVIAPKDRVYEKILSNV 517
QY 535 BEVRARGQLVVFADQDAGFVSSDNMHIEMPHVEEVIAPIFYVPLQLLAYHVALIKGT 594
Db 518 BEVLARKGRVISVGKGDTELKSKSESVMIBPKAEBIPFPLTVIPLQLFAYFIASKLGL 577
QY 595 DVDQPRNLAKSVTVE 609
Db 578 DVDQPRNLAKTVTVE 592
RESULT 15
US-09-069-226-30
; Sequence 30, Application US/09069226
; Patent No. 6013509
; GENERAL INFORMATION:
; APPLICANT: WARREN, Patrick V.
; TITLE OF INVENTION: TRANSAMINASES AND AMINOTRANSFERASES
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/069,226
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/599,171
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: HERRON, CHARLES J.
; REGISTRATION NUMBER: 28,019
; REFERENCE/DOCKET NUMBER: 331400-38
; TELECOMMUNICATION INFORMATION:


```
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 592 AMINO ACIDS
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
US-09-069-226-30

Query Match      44.0%; Score 1357.5; DB 2; Length 592;
Best Local Similarity 46.5%; Pred. No. 1.3e-124;
Matches 286; Conservative 113; Mismatches 187; Indels 29; Gaps 9;

Qy      1 MCGIVGAIAORDVA-EILLEGRLREYRGYDSAGLAVVDAEGHMTLRLRLGKVMQLAQAA 59
Db      1 MCGIVGVVG-RDLALPIVLGALERLEYRGYDSAGVALIE-DGKLIVVEKKKGKRELKAL 58

Qy      60 EEHPLHGCTGIAHTRWATHGEPSEVNAHPHVE--HIVVVHNGIIEHNHEPLREELKARG 117
Db      59 WGDYKAKTGIGTRWATHGKPTDENAHPTDEKGEFAVVHNGIIEHNYLELKEELKEGV 118

Qy      118 TFVSETDTEVIAHLVNWELKQGTLEAVLRAPQLRGAYGTVMDSRHPDPTLLAARSGS 177
Db      119 KFRSETDTEVIAHLIAKNYR--GDLEAVLTKVKLKGAFATVITVHEPNRLIGVKQGS 176

Qy      178 PLVIGLGMGNFIASDQALLPVTTRRFLEEGDIAEITRRSVNIPDKTGAEVKRODIES 237
Db      177 PLIVGLGEGENFLASDIPALPYTKIIVLDGGEIADLTPTDVTNINFEGEPSKVEWIT 236

Qy      238 NLQYDAGDKGIYRHYMQEIEYQPNAIKNTLTGRISHGQVDLSLGPNADEL---LSKVE 294
Db      237 PWDIVSAEKGFRHFMLKEIYEQPKAINDTLKGFLS-----TEDAIPFKLKDFR 285

Qy      295 HIQILACGTSYNSGMVRYWPFESLAGIPCDVEIASEPRYRKSAVRNLSMITLSQGETA 354
Db      286 RVLIIACGTSYHAGFVGKYWIERFAGVPTFVIYASEFRYADVPVSKDIDIVIGISQGETA 345

Qy      355 DTLAGRLSKELGYLSGLAICNVPGSSLVRESDLALMTNAGTEIGVASTKRAFTOLTVL 414
Db      346 DTFALQSKEKGAF-TVGLNVNVVGSARDRESDFSLTHAGPEIGVAATKTFTAQFTALY 404

Qy      415 MLVAKLSRLKGLDASIEHDIVHGIQALPFSRIEQMLSDKRIEALAEFSDKHHALFLGRG 474
Db      405 ALSVR-----ESEERENLIRLLEKVPVSLVEQTLNTAEVEKVAEKYMKKNWLYLGRY 457

Qy      475 DQYPIALEGALKLKEISYTHAEAYAGELKHGPIALIDAMPVTVVAPNNELLEKLKSN 534
Db      458 LNTYPIALEGALKLKEISYTHAEGYPAGEMKHGPIALIDENMPVVVVIAPKDRVYEKILSNV 517

Qy      535 EEVPRAGQLYVFADQDAGFVSSDNMHIIEHPHVEEVIPIFYTVPLQLLAYHVALIKGT 594
Db      518 EEVLARKGRVISVGFGDETLKSKSVMEIPKAEPEITPFLTVPILQLFAYFIASKLGL 577

Qy      595 DVDQPRNLAKSVTVE 609
Db      578 DVDQPRNLAKTVTVE 592
```

Search completed: August 7, 2006, 09:26:25
Job time : 26 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 7, 2006, 09:26:41 ; Search time 93 Seconds
(without alignments)
3033.310 Million cell updates/sec

Title: US-10-612-779-2

Perfect score: 3086

Sequence: 1 MCGIVGAIQRDVAEILLEG.....LIKGTDVDPRLNAKSVTVE 609

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA Main:*

- 1: /EMC_Celerra_SID83/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /EMC_Celerra_SID83/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 3: /EMC_Celerra_SID83/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 4: /EMC_Celerra_SID83/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
- 5: /EMC_Celerra_SID83/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 6: /EMC_Celerra_SID83/ptodata/2/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|--|
| 1 | 3086 | 100.0 | 609 | 3 | US-09-815-242-10399 Sequence 10399, A |
| 2 | 3086 | 100.0 | 609 | 4 | US-10-024-460-16 Sequence 16, Appl |
| 3 | 3086 | 100.0 | 609 | 4 | US-10-369-493-23611 Sequence 23611, A |
| 4 | 3086 | 100.0 | 609 | 4 | US-10-282-122A-56753 Sequence 56753, A |
| 5 | 3086 | 100.0 | 609 | 4 | US-10-612-779-2 Sequence 2, Appl |
| 6 | 3086 | 100.0 | 609 | 5 | US-10-823-397-23 Sequence 23, Appl |
| 7 | 3080 | 99.8 | 609 | 4 | US-10-024-460-28 Sequence 28, Appl |
| 8 | 3080 | 99.8 | 609 | 4 | US-10-024-460-31 Sequence 31, Appl |
| 9 | 3080 | 99.8 | 609 | 4 | US-10-612-779-10 Sequence 10, Appl |
| 10 | 3080 | 99.8 | 609 | 4 | US-10-612-779-12 Sequence 12, Appl |
| 11 | 3079 | 99.8 | 609 | 4 | US-10-024-460-25 Sequence 25, Appl |
| 12 | 3079 | 99.8 | 609 | 4 | US-10-612-779-8 Sequence 8, Appl |
| 13 | 3071 | 99.5 | 609 | 4 | US-10-024-460-19 Sequence 19, Appl |
| 14 | 3071 | 99.5 | 609 | 4 | US-10-612-779-4 Sequence 4, Appl |
| 15 | 3068 | 99.4 | 609 | 4 | US-10-024-460-22 Sequence 22, Appl |
| 16 | 3068 | 99.4 | 609 | 4 | US-10-612-779-6 Sequence 6, Appl |
| 17 | 3068 | 99.4 | 609 | 5 | US-10-823-397-25 Sequence 25, Appl |
| 18 | 3061 | 99.2 | 609 | 4 | US-10-612-779-14 Sequence 14, Appl |
| 19 | 3053 | 98.9 | 609 | 3 | US-09-815-242-14065 Sequence 14065, A |
| 20 | 3053 | 98.9 | 609 | 4 | US-10-282-122A-76152 Sequence 76152, A |
| 21 | 2950 | 95.6 | 609 | 4 | US-10-282-122A-56279 Sequence 56279, A |
| 22 | 2876 | 93.2 | 609 | 3 | US-09-815-242-11758 Sequence 11758, A |
| 23 | 2871 | 93.0 | 609 | 4 | US-10-282-122A-59335 Sequence 59335, A |
| 24 | 2718.5 | 88.1 | 608 | 4 | US-10-282-122A-72965 Sequence 72965, A |
| 25 | 2682 | 86.9 | 609 | 4 | US-10-282-122A-78533 Sequence 78533, A |
| 26 | 2661 | 86.2 | 609 | 4 | US-10-369-493-21242 Sequence 21242, A |
| 27 | 2619 | 84.9 | 609 | 4 | US-10-369-493-438 Sequence 438, App |

ALIGNMENTS

RESULT 1

US-09-815-242-10399
; Sequence 10399, Application US/09815242
; Patent No. US20020061569A1

GENERAL INFORMATION:

- ; APPLICANT: Haseelbeck, Robert
- ; APPLICANT: Ohlsen, Kari L.
- ; APPLICANT: Zyskind, Judith W.
- ; APPLICANT: Wall, Daniel
- ; APPLICANT: Trawick, John D.
- ; APPLICANT: Carr, Grant J.
- ; APPLICANT: Yamamoto, Robert T.
- ; APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in

FILE OF INVENTION: Prokaryotes

FILE REFERENCE: ELITPA.011A

CURRENT APPLICATION NUMBER: US/09/815,242

CURRENT FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 10399

LENGTH: 609

TYPE: PRT

ORGANISM: Escherichia coli

US-09-815-242-10399

Query Match 100.0%; Score 3086; DB 3; Length 609;

Best Local Similarity 100.0%; Pred. No. 4.7e-249;

Matches 609; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MCGIVGAIQRDVAEILLEGRLREYRGYSAGLVAVDAEGHMTLRLKQVQMLAAAE 60

Db 1 MCGIVGAIQRDVAEILLEGRLREYRGYSAGLVAVDAEGHMTLRLKQVQMLAAAE 60

Qy 61 EPHLHGDTGTAHTRWATHGSEPNAPHYSEHVIVVHNGIENHEPLRELKARGYTFV 120

Db 61 EHPLHGGTGAHTRWATHGEPSEVNAHPHVSEHIVVHNGIIENHEPLREELKARGYTFV 120
QY 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDPTLLAARSGSPLV 180
Db 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDPTLLAARSGSPLV 180
QY 181 IGLGNGENFIASDQALALLPVTRRFIFLEEGDIAETIRRSVNIIPDKTGAEVKRODIESNLQ 240
Db 181 IGLGNGENFIASDQALALLPVTRRFIFLEEGDIAETIRRSVNIIPDKTGAEVKRODIESNLQ 240
QY 241 YDAGDKGIYRHYMQKEIYEQPNAINKNTLTGRISHGQVDLSSELGNADDELLSKVEHIQILA 300
Db 241 YDAGDKGIYRHYMQKEIYEQPNAINKNTLTGRISHGQVDLSSELGNADDELLSKVEHIQILA 300
QY 301 CGTSYNSGMVSRWYFESLAGIPCDVEIASFEFRYKSAVRNSLMITLSQSGETADTLAGL 360
Db 301 CGTSYNSGMVSRWYFESLAGIPCDVEIASFEFRYKSAVRNSLMITLSQSGETADTLAGL 360
QY 361 RLSKELGYLGSALAINCVPGSSLVRESDLALMTNAGTEIGVASTKAFPTTQTLTVLLMLVAKL 420
Db 361 RLSKELGYLGSALAINCVPGSSLVRESDLALMTNAGTEIGVASTKAFPTTQTLTVLLMLVAKL 420
QY 421 SRLKGLDASIEHDIVHGLQALPSRIEQMLSDQKRIEALAEFSDKHAFFLGRGDQYPIA 480
Db 421 SRLKGLDASIEHDIVHGLQALPSRIEQMLSDQKRIEALAEFSDKHAFFLGRGDQYPIA 480
QY 481 LEGALKLKEISYTHAEAYAAGELKHGPIALIDAMPVIVVAPNNELLEKLKSNIEBVRAR 540
Db 481 LEGALKLKEISYTHAEAYAAGELKHGPIALIDAMPVIVVAPNNELLEKLKSNIEBVRAR 540
QY 541 GGQLYVPADQAGFVSSDNNHIIEMPHVEEVIAPIFTYVPLQLLAYHVALIKGTDVDDQR 600
Db 541 GGQLYVPADQAGFVSSDNNHIIEMPHVEEVIAPIFTYVPLQLLAYHVALIKGTDVDDQR 600
QY 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 2

US-10-024-460-16
; Sequence 16, Application US/10024460
; Publication No. US20030044939A1
; GENERAL INFORMATION:
; APPLICANT: Berry, Alan
; APPLICANT: Burlingame, Richard P.
; APPLICANT: Millis, James R.
; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE
; FILE REFERENCE: 3161-18-C1
; CURRENT APPLICATION NUMBER: US/10/024,460
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/115,475
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/035,494
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-01-14
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-024-460-16

Query Match 100.0%; Score 3086; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 4.7e-249;
Matches 609; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MCGIVGAIAQRDVAEIILLEGRLREYRGYDSAGLAVDAEGHMTLRRLRGKQVQLAQAAE 60
Db 1 MCGIVGAIAQRDVAEIILLEGRLREYRGYDSAGLAVDAEGHMTLRRLRGKQVQLAQAAE 60
QY 61 EHPLHGGTGAHTRWATHGEPSEVNAHPHVSEHIVVHNGIIENHEPLREELKARGYTFV 120

Db 61 EHPLHGGTGAHTRWATHGEPSEVNAHPHVSEHIVVHNGIIENHEPLREELKARGYTFV 120
QY 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDPTLLAARSGSPLV 180
Db 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPOLRGAYGTVMDSRHPDPTLLAARSGSPLV 180
QY 181 IGLGNGENFIASDQALALLPVTRRFIFLEEGDIAETIRRSVNIIPDKTGAEVKRODIESNLQ 240
Db 181 IGLGNGENFIASDQALALLPVTRRFIFLEEGDIAETIRRSVNIIPDKTGAEVKRODIESNLQ 240
QY 241 YDAGDKGIYRHYMQKEIYEQPNAINKNTLTGRISHGQVDLSSELGNADDELLSKVEHIQILA 300
Db 241 YDAGDKGIYRHYMQKEIYEQPNAINKNTLTGRISHGQVDLSSELGNADDELLSKVEHIQILA 300
QY 301 CGTSYNSGMVSRWYFESLAGIPCDVEIASFEFRYKSAVRNSLMITLSQSGETADTLAGL 360
Db 301 CGTSYNSGMVSRWYFESLAGIPCDVEIASFEFRYKSAVRNSLMITLSQSGETADTLAGL 360
QY 361 RLSKELGYLGSALAINCVPGSSLVRESDLALMTNAGTEIGVASTKAFPTTQTLTVLLMLVAKL 420
Db 361 RLSKELGYLGSALAINCVPGSSLVRESDLALMTNAGTEIGVASTKAFPTTQTLTVLLMLVAKL 420
QY 421 SRLKGLDASIEHDIVHGLQALPSRIEQMLSDQKRIEALAEFSDKHAFFLGRGDQYPIA 480
Db 421 SRLKGLDASIEHDIVHGLQALPSRIEQMLSDQKRIEALAEFSDKHAFFLGRGDQYPIA 480
QY 481 LEGALKLKEISYTHAEAYAAGELKHGPIALIDAMPVIVVAPNNELLEKLKSNIEBVRAR 540
Db 481 LEGALKLKEISYTHAEAYAAGELKHGPIALIDAMPVIVVAPNNELLEKLKSNIEBVRAR 540
QY 541 GGQLYVPADQAGFVSSDNNHIIEMPHVEEVIAPIFTYVPLQLLAYHVALIKGTDVDDQR 600
Db 541 GGQLYVPADQAGFVSSDNNHIIEMPHVEEVIAPIFTYVPLQLLAYHVALIKGTDVDDQR 600
QY 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 3

US-10-369-493-23611
; Sequence 23611, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 23611
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-369-493-23611

Query Match 100.0%; Score 3086; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 4.7e-249;
Matches 609; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MCGIVGAIAQRDVAEIILLEGRLREYRGYDSAGLAVDAEGHMTLRRLRGKQVQLAQAAE 60
Db 1 MCGIVGAIAQRDVAEIILLEGRLREYRGYDSAGLAVDAEGHMTLRRLRGKQVQLAQAAE 60
QY 61 EHPLHGGTGAHTRWATHGEPSEVNAHPHVSEHIVVHNGIIENHEPLREELKARGYTFV 120

Db 61 EHPHGGGTGIAHTRWATHGEPSEVNAHPHVSEHIVVHHNGIIENHPELREELKARGYTFV 120
Qy 121 SETDTEVIAHLVNWELKQGGLTLEAVLRAIPQURGAYGTVMDSRHPDPTLLAARSGSPLV 180
Db 121 SETDTEVIAHLVNWELKQGGLTLEAVLRAIPQURGAYGTVMDSRHPDPTLLAARSGSPLV 180
Qy 181 IGLGMENFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIIPDKTGAEVKRDIESNLQ 240
Db 181 IGLGMENFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIIPDKTGAEVKRDIESNLQ 240
Qy 241 YDAGDKGIYRHYMQKEIYEOPNAIKNTLTGRISHGQVDLSSELGNADDELLSKVEHIQILA 300
Db 241 YDAGDKGIYRHYMQKEIYEOPNAIKNTLTGRISHGQVDLSSELGNADDELLSKVEHIQILA 300
Qy 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASFPRYKSAVRNLSMITLSQSGETADTLA 360
Db 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASFPRYKSAVRNLSMITLSQSGETADTLA 360
Qy 361 RLSKELGYLGSGLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFPTTQLTVLLMLVAKL 420
Db 361 RLSKELGYLGSGLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFPTTQLTVLLMLVAKL 420
Qy 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMLSQDKRIEALAEFSDKHHALFLGRGDQYPIA 480
Db 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMLSQDKRIEALAEFSDKHHALFLGRGDQYPIA 480
Qy 481 LEGALKLKEISYTHAEAYAGELKHGPLALIDAMPVIVVAPNNELLEKLKSNIEVRAR 540
Db 481 LEGALKLKEISYTHAEAYAGELKHGPLALIDAMPVIVVAPNNELLEKLKSNIEVRAR 540
Qy 541 GGOLYVFADQDAGFVSSDNNHIIEMPHVEEVIAPIFTVPLQLLAYHVALIKGTDVDDQPR 600
Db 541 GGOLYVFADQDAGFVSSDNNHIIEMPHVEEVIAPIFTVPLQLLAYHVALIKGTDVDDQPR 600
Qy 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 4
US-10-282-122A-56753
; Sequence 56753, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zvekand, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.0344
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 56753
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-10-282-122A-56753

Query Match 100.0%; Score 3086; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 4,7e-249;
Matches 609; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MCGIVGAIQAORDVAEILLEGRLREYRGYDSAGLAVVDAEGHMTLRLRLGKVQMLAQAAE 60
Db 1 MCGIVGAIQAORDVAEILLEGRLREYRGYDSAGLAVVDAEGHMTLRLRLGKVQMLAQAAE 60
Qy 61 EHPHGGGTGIAHTRWATHGEPSEVNAHPHVSEHIVVHHNGIIENHPELREELKARGYTFV 120
Db 61 EHPHGGGTGIAHTRWATHGEPSEVNAHPHVSEHIVVHHNGIIENHPELREELKARGYTFV 120
Qy 121 SETDTEVIAHLVNWELKQGGLTLEAVLRAIPQURGAYGTVMDSRHPDPTLLAARSGSPLV 180
Db 121 SETDTEVIAHLVNWELKQGGLTLEAVLRAIPQURGAYGTVMDSRHPDPTLLAARSGSPLV 180
Qy 181 IGLGMENFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIIPDKTGAEVKRDIESNLQ 240
Db 181 IGLGMENFIASDQALLPVTRRFIFLEEGDIAEITRRSVNIIPDKTGAEVKRDIESNLQ 240
Qy 241 YDAGDKGIYRHYMQKEIYEOPNAIKNTLTGRISHGQVDLSSELGNADDELLSKVEHIQILA 300
Db 241 YDAGDKGIYRHYMQKEIYEOPNAIKNTLTGRISHGQVDLSSELGNADDELLSKVEHIQILA 300
Qy 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASFPRYKSAVRNLSMITLSQSGETADTLA 360
Db 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASFPRYKSAVRNLSMITLSQSGETADTLA 360
Qy 361 RLSKELGYLGSGLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFPTTQLTVLLMLVAKL 420
Db 361 RLSKELGYLGSGLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFPTTQLTVLLMLVAKL 420
Qy 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMLSQDKRIEALAEFSDKHHALFLGRGDQYPIA 480
Db 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMLSQDKRIEALAEFSDKHHALFLGRGDQYPIA 480
Qy 481 LEGALKLKEISYTHAEAYAGELKHGPLALIDAMPVIVVAPNNELLEKLKSNIEVRAR 540
Db 481 LEGALKLKEISYTHAEAYAGELKHGPLALIDAMPVIVVAPNNELLEKLKSNIEVRAR 540
Qy 541 GGOLYVFADQDAGFVSSDNNHIIEMPHVEEVIAPIFTVPLQLLAYHVALIKGTDVDDQPR 600
Db 541 GGOLYVFADQDAGFVSSDNNHIIEMPHVEEVIAPIFTVPLQLLAYHVALIKGTDVDDQPR 600
Qy 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 5
US-10-612-779-2
; Sequence 2, Application US/10612779
; Publication No. US20040091976A1
; GENERAL INFORMATION:
; APPLICANT: Deng, Ming-De
; APPLICANT: Angerer, J. David
; APPLICANT: Cyron, Don

APPLICANT: Grund, Alan
APPLICANT: Jerrell Jr., Thomas
APPLICANT: Leanna, Candice
APPLICANT: Mathre, Owen
APPLICANT: Robson, Reinhardt
APPLICANT: Running, Jeff
APPLICANT: Severson, Dave
APPLICANT: Song, Linsheng
APPLICANT: Wasink, Sarah
TITLE OF INVENTION: Process and Materials for Production of Glucosamine and N-Acetylg
TITLE OF INVENTION: Glucosamine
FILE REFERENCE: 3161-18-2
CURRENT APPLICATION NUMBER: US/10/612,779
CURRENT FILING DATE: 2003-07-01
PRIOR APPLICATION NUMBER: 60/393,348
PRIOR FILING DATE: 2002-07-01
NUMBER OF SEQ ID NOS: 137
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 609
TYPE: PRT
ORGANISM: Escherichia coli
US-10-612-779-2

Query Match 100.0%; Score 3086; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 4.7e-249;
Matches 609; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 MCGIVGAIARDVNAEILLEGRLRLRYGYSAGLAVVDAEGHMTLRRLRGKQVQLAAQAE 60
Db 1 MCGIVGAIARDVNAEILLEGRLRLRYGYSAGLAVVDAEGHMTLRRLRGKQVQLAAQAE 60
Qy 61 EPHLHGTTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEPLREELKARGYTFV 120
Db 61 EPHLHGTTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEPLREELKARGYTFV 120
Qy 121 SETDTEVIAHLVNMELKQGGTLREAVLRAIPQLRGAGYTVIMDSRHPDTLLAARSGSPV 180
Db 121 SETDTEVIAHLVNMELKQGGTLREAVLRAIPQLRGAGYTVIMDSRHPDTLLAARSGSPV 180
Qy 181 IGLGMENFTASDQALLPVTRRFIFLEEGDIAEITRRSVNIPDKTGAEVKQDIESNLQ 240
Db 181 IGLGMENFTASDQALLPVTRRFIFLEEGDIAEITRRSVNIPDKTGAEVKQDIESNLQ 240
Qy 241 YDAGDKGIYHYMKEIYEOPNAIKNTLTGRISHGOVDLSLGPNADELLSKVEHIQILA 300
Db 241 YDAGDKGIYHYMKEIYEOPNAIKNTLTGRISHGOVDLSLGPNADELLSKVEHIQILA 300
Qy 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASEFRYRKSAVRRNSLMTLSQSGETADTLA 360
Db 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASEFRYRKSAVRRNSLMTLSQSGETADTLA 360
Qy 361 RLSKELGYLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAPTTLTLLMLVAKL 420
Db 361 RLSKELGYLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAPTTLTLLMLVAKL 420
Qy 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMLSDQKRIEALAEFSDKHALLFLGRGDQYPTA 480
Db 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMLSDQKRIEALAEFSDKHALLFLGRGDQYPTA 480
Qy 481 LEGALKLKEISYTHAEAYAAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540
Db 481 LEGALKLKEISYTHAEAYAAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540
Qy 541 GGQLYVFADQDAGFVSSDNHIIEMPHVEEVIAPFYTVPLQLLAYHVALIKGTDVDDPR 600
Db 541 GGQLYVFADQDAGFVSSDNHIIEMPHVEEVIAPFYTVPLQLLAYHVALIKGTDVDDPR 600
Qy 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609
```

RESULT 7

RESULT 6
US-10-823-397-23
Sequence 23, Application US/10823397
Publication No. US20050042735A1
GENERAL INFORMATION:
APPLICANT: McMullin, Thomas
APPLICANT: Ding, Ming-De
APPLICANT: Grund, Alan
TITLE OF INVENTION: Metabolic Engineering for Enhanced Production of Chitin and
TITLE OF INVENTION: Chitosan in Microorganisms
FILE REFERENCE: 3161-18-3
CURRENT APPLICATION NUMBER: US/10/823,397
CURRENT FILING DATE: 2004-04-12
PRIOR APPLICATION NUMBER: 60/462,087
PRIOR FILING DATE: 2003-04-11
NUMBER OF SEQ ID NOS: 59
SOFTWARE: PatentIn version 3.2
SEQ ID NO 23
LENGTH: 609
TYPE: PRT
ORGANISM: Escherichia coli
US-10-823-397-23

```
Query Match 100.0%; Score 3086; DB 5; Length 609;  
Best Local Similarity 100.0%; Pred. No. 4.7e-249;  
Matches 609; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 MCGIVGAIARDVNAEILLEGRLRLRYGYSAGLAVVDAEGHMTLRRLRGKQVQLAAQAE 60
Db 1 MCGIVGAIARDVNAEILLEGRLRLRYGYSAGLAVVDAEGHMTLRRLRGKQVQLAAQAE 60
Qy 61 EPHLHGTTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEPLREELKARGYTFV 120
Db 61 EPHLHGTTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEPLREELKARGYTFV 120
Qy 121 SETDTEVIAHLVNMELKQGGTLREAVLRAIPQLRGAGYTVIMDSRHPDTLLAARSGSPV 180
Db 121 SETDTEVIAHLVNMELKQGGTLREAVLRAIPQLRGAGYTVIMDSRHPDTLLAARSGSPV 180
Qy 181 IGLGMENFTASDQALLPVTRRFIFLEEGDIAEITRRSVNIPDKTGAEVKQDIESNLQ 240
Db 181 IGLGMENFTASDQALLPVTRRFIFLEEGDIAEITRRSVNIPDKTGAEVKQDIESNLQ 240
Qy 241 YDAGDKGIYHYMKEIYEOPNAIKNTLTGRISHGOVDLSLGPNADELLSKVEHIQILA 300
Db 241 YDAGDKGIYHYMKEIYEOPNAIKNTLTGRISHGOVDLSLGPNADELLSKVEHIQILA 300
Qy 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASEFRYRKSAVRRNSLMTLSQSGETADTLA 360
Db 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASEFRYRKSAVRRNSLMTLSQSGETADTLA 360
Qy 361 RLSKELGYLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAPTTLTLLMLVAKL 420
Db 361 RLSKELGYLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAPTTLTLLMLVAKL 420
Qy 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMLSDQKRIEALAEFSDKHALLFLGRGDQYPTA 480
Db 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMLSDQKRIEALAEFSDKHALLFLGRGDQYPTA 480
Qy 481 LEGALKLKEISYTHAEAYAAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540
Db 481 LEGALKLKEISYTHAEAYAAAGELKHGPLALIDADMPVIVVAPNNELLEKLKSNIEVRAR 540
Qy 541 GGQLYVFADQDAGFVSSDNHIIEMPHVEEVIAPFYTVPLQLLAYHVALIKGTDVDDPR 600
Db 541 GGQLYVFADQDAGFVSSDNHIIEMPHVEEVIAPFYTVPLQLLAYHVALIKGTDVDDPR 600
Qy 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609
```

```

US-10-024-460-28
; Sequence 28, Application US/10024460
; Publication No. US20030044939A1
; GENERAL INFORMATION:
; APPLICANT: Berry, Alan
; APPLICANT: Burlingame, Richard P.
; APPLICANT: Millis, James R.
; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE
; FILE REFERENCE: 3161-18-C1
; CURRENT APPLICATION NUMBER: US/10/024,460
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/115,475
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/035,494
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-01-14
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-024-460-28

```

RESULT 8

```

US-10-024-460-31
; Sequence 31, Application US/10024460
; Publication No. US20030044939A1
; GENERAL INFORMATION:
; APPLICANT: Berry, Alan
; APPLICANT: Burlingame, Richard P.
; APPLICANT: Mullis, James R.
; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE
; FILE REFERENCE: 3161-18-C1
; CURRENT APPLICATION NUMBER: US/10/024,460
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/115,475
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/035,494
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-01-14
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31
; LENGTH: 609
; TYPE: PRF
; ORGANISM: Escherichia coli
US-10-024-460-31

```

RESULT 9

US-10-612-779-10
; Sequence 10, Application US/10612779
; Publication NO. US20040091976A1
; GENERAL INFORMATION:
; APPLICANT: Deng, Ming-De
; APPLICANT: Angerer, J. David
; APPLICANT: Cyron, Don
; APPLICANT: Grund, Alan
; APPLICANT: Jerrell Jr., Thomas
; APPLICANT: Leanna, Candice
; APPLICANT: Mathre, Owen
; APPLICANT: Rosson, Reinhardt
; APPLICANT: Running, Jeff
; APPLICANT: Severson, Dave
; APPLICANT: Song, Linsheng
; APPLICANT: Wassink, Sarah
; TITLE OF INVENTION: Process and Materials for Production of Glucosamine and N-Acetyls
; FILE REFERENCE: 3161-18-2
; CURRENT APPLICATION NUMBER: US/10/612,779
; CURRENT FILING DATE: 2003-07-01
; PRIOR APPLICATION NUMBER: 60/393,348
; PRIOR FILING DATE: 2002-07-01
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-612-779-10

Query Match 99.8%; Score 3080; DB 4; Length 609;
Best Local Similarity 99.8%; Pred. No. 1.5e-248;
Matches 608; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MCGIVGATGAQDVAAEILLEGRLRLEYRGYDSAGLAVVDAEGHMTLRRLRGKQVQLAAQAE 60
DB 1 MCGIVGATGAQDVAAEILLEGRLRLEYRGYDSAGLAVVDAEGHMTLRRLRGKQVQLAAQAE 60

QY 61 EPHLGGTGTIAHTRWATHGEPSEVNAPHVSEHIVVHNGIIEHNEPLREELKARGYTFV 120
DB 61 EPHLGGTGTIAHTRWATHGEPSEVNAPHVSEHIVVHNGIIEHNEPLREELKARGYTFV 120

QY 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDPTLLAARSGSPLV 180
DB 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDPTLLAARSGSPLV 180

QY 181 IGLGMENFIASDQALLPVTRRFIFLEEGDIAETRRSVNIFDKTGAEVKQDIESNLQ 240
DB 181 IGLGMENFIASDQALLPVTRRFIFLEEGDIAETRRSVNIFDKTGAEVKQDIESNLQ 240

QY 241 YDAGDKGIYRHYMQKEIYEOPNAIKNTLTGRISHGQVDLSLGNADLLSKVEHIQILA 300
DB 241 YDAGDKGIYRHYMQKEIYEOPNAIKNTLTGRISHGQVDLSLGNADLLSKVEHIQILA 300

QY 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASFPRYKSAVRNSIMITLSQSGETADTLAGL 360
DB 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASFPRYKSAVRNSIMITLSQSGETADTLAGL 360

QY 361 RLSKELGYLGSIALCNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQTLTVLLMLVAKL 420
DB 361 RLSKELGYLGSIALCNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQTLTVLLMLVAKL 420

QY 421 SRLKGLDASIEHDIHVHGLQALPSRIEOMLSQDKRIEALAEFSDKHAFILGRGQDYPTA 480
DB 421 SRLKGLDASIEHDIHVHGLQALPSRIEOMLSQDKRIEALAEFSDKHAFILGRGQDYPTA 480

QY 481 LEGALKKEISYTHAEYAGELKHGFLALIDADMPVIVVAPNNLEKLSNIEVRAR 540
DB 481 LEGALKKEISYTHAEYAGELKHGFLALIDADMPVIVVAPNNLEKLSNIEVRAR 540

QY 541 GGQLYVFAQQDAGFVSSDNNHIIEMPHVEVIAPIFYTVPLQLLAVHVALIKGTDVQDPR 600
DB 541 GGQLYVFAQQDAGFVSSDNNHIIEMPHVEVIAPIFYTVPLQLLAVHVALIKGTDVQDPR 600

DB 541 GGQLYVFAQQDAGFVSSDNNHIIEMPHVEVIAPIFYTVPLQLLAVHVALIKGTDVQDPR 600
QY 601 NLAKSVTVE 609
DB 601 NLAKSVTVE 609

RESULT 10
US-10-612-779-12
; Sequence 12, Application US/10612779
; Publication NO. US20040091976A1
; GENERAL INFORMATION:
; APPLICANT: Deng, Ming-De
; APPLICANT: Angerer, J. David
; APPLICANT: Cyron, Don
; APPLICANT: Grund, Alan
; APPLICANT: Jerrell Jr., Thomas
; APPLICANT: Leanna, Candice
; APPLICANT: Mathre, Owen
; APPLICANT: Rosson, Reinhardt
; APPLICANT: Running, Jeff
; APPLICANT: Severson, Dave
; APPLICANT: Song, Linsheng
; APPLICANT: Wassink, Sarah
; TITLE OF INVENTION: Process and Materials for Production of Glucosamine and N-Acetyls
; FILE REFERENCE: 3161-18-2
; CURRENT APPLICATION NUMBER: US/10/612,779
; CURRENT FILING DATE: 2003-07-01
; PRIOR APPLICATION NUMBER: 60/393,348
; PRIOR FILING DATE: 2002-07-01
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-612-779-12

Query Match 99.8%; Score 3080; DB 4; Length 609;
Best Local Similarity 99.8%; Pred. No. 1.5e-248;
Matches 608; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MCGIVGATGAQDVAAEILLEGRLRLEYRGYDSAGLAVVDAEGHMTLRRLRGKQVQLAAQAE 60
DB 1 MCGIVGATGAQDVAAEILLEGRLRLEYRGYDSAGLAVVDAEGHMTLRRLRGKQVQLAAQAE 60

QY 61 EPHLGGTGTIAHTRWATHGEPSEVNAPHVSEHIVVHNGIIEHNEPLREELKARGYTFV 120
DB 61 EPHLGGTGTIAHTRWATHGEPSEVNAPHVSEHIVVHNGIIEHNEPLREELKARGYTFV 120

QY 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDPTLLAARSGSPLV 180
DB 121 SETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDPTLLAARSGSPLV 180

QY 181 IGLGMENFIASDQALLPVTRRFIFLEEGDIAETRRSVNIFDKTGAEVKQDIESNLQ 240
DB 181 IGLGMENFIASDQALLPVTRRFIFLEEGDIAETRRSVNIFDKTGAEVKQDIESNLQ 240

QY 241 YDAGDKGIYRHYMQKEIYEOPNAIKNTLTGRISHGQVDLSLGNADLLSKVEHIQILA 300
DB 241 YDAGDKGIYRHYMQKEIYEOPNAIKNTLTGRISHGQVDLSLGNADLLSKVEHIQILA 300

QY 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASFPRYKSAVRNSIMITLSQSGETADTLAGL 360
DB 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASFPRYKSAVRNSIMITLSQSGETADTLAGL 360

QY 361 RLSKELGYLGSIALCNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQTLTVLLMLVAKL 420
DB 361 RLSKELGYLGSIALCNVPGSSLVRESDLALMTNAGTEIGVASTKFTTQTLTVLLMLVAKL 420

QY 421 SRLKGLDASIEHDIHVHGLQALPSRIEOMLSQDKRIEALAEFSDKHAFILGRGQDYPTA 480
DB 421 SRLKGLDASIEHDIHVHGLQALPSRIEOMLSQDKRIEALAEFSDKHAFILGRGQDYPTA 480

Db 421 SRLKGLDASIEHDIHVHGLQALPSRIEQLSODKRIEALAEDFSQKHALFLSRGQVPIA 480
Qy 481 LEGALKLKEISYIHAEYAAAGELKHGPLALIDAMPVIVVAPNNLELEKLSNIEEVRAR 540
Db 481 LEGALKLKEISYIHAEYAAAGELKHGPLALIDAMPVIVVAPNNLELEKLSNIEEVRAR 540
Qy 541 GGQLYVPADQAGFVSSDNMHHIEMPHVEEVIAPIFTYVPLQLLAYHVALIKGTDVQDPR 600
Db 541 GGQLYVPADQAGFVSSDNMHHIEMPHVEEVIAPIFTYVPLQLLAYHVALIKGTDVQDPR 600
Qy 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609
RESULT 11
US-10-024-460-25
; Sequence 25, Application US/10024460
; Publication No. US20030044939A1
; GENERAL INFORMATION:
; APPLICANT: Berry, Alan
; APPLICANT: Burlingame, Richard P.
; APPLICANT: Millis, James R.
; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE
; FILE REFERENCE: 3161-18-C1
; CURRENT APPLICATION NUMBER: US/10/024,460
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/115,475
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/035,494
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-01-14
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-024-460-25
Query Match 99.8%; Score 3079; DB 4; Length 609;
Best Local Similarity 99.8%; Pred. No. 1.8e-248;
Matches 608; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MCGIVGAIAQORDVAEILLEGRLRLLEYRGYDSAGLAVVDAGHMTLRRLRGKVQMLAQAAE 60
Db 1 MCGIVGAIAQORDVAEILLEGRLRLLEYRGYDSAGLAVVDAGHMTLRRLRGKVQMLAQAAE 60
Qy 61 EPHLHGCTGIAHTRWATHGEPSEVNAHPHVSEHIVVHNGIENHPELREBELKARGYTFV 120
Db 61 EPHLHGCTGIAHTRWATHGEPSEVNAHPHVSEHIVVHNGIENHPELREBELKARGYTFV 120
Qy 121 SETDTEVIAHLVNWELKQGGLTREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPLV 180
Db 121 SETDTEVIAHLVNWELKQGGLTREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPLV 180
Qy 181 IGLGMGNFTASDQALLPVTRRFIFLEEGDIAETRRSVNIFDKTGAEVKRODIESNLQ 240
Db 181 IGLGMGNFTASDQALLPVTRRFIFLEEGDIAETRRSVNIFDKTGAEVKRODIESNLQ 240
Qy 241 YDAGDKGIYRHYMQKEIYEOPNAIKNTLTGRISHGQVDLSELGPNADLLSKVEHIQILA 300
Db 241 YDAGDKGIYRHYMQKEIYEOPNAIKNTLTGRISHGQVDLSELGPNADLLSKVEHIQILA 300
Qy 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASFPRYKSAVRNSLMITLQSGETADTLAGL 360
Db 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASFPRYKSAVRNSLMITLQSGETADTLAGL 360
Qy 361 RLSKELGYLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFQTTLQTLVLLMLVAKL 420
Db 361 RLSKELGYLSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFQTTLQTLVLLMLVAKL 420
Qy 421 SRLKGLDASIEHDIHVHGLQALPSRIEQLSODKRIEALAEDFSQKHALFLSRGQVPIA 480

Db 421 SRLKGLDASIEHDIHVHGLQALPSRIEQLSODKRIEALAEDFSQKHALFLSRGQVPIA 480
Qy 481 LEGALKLKEISYIHAEYAAAGELKHGPLALIDAMPVIVVAPNNLELEKLSNIEEVRAR 540
Db 481 LEGALKLKEISYIHAEYAAAGELKHGPLALIDAMPVIVVAPNNLELEKLSNIEEVRAR 540
Qy 541 GGQLYVPADQAGFVSSDNMHHIEMPHVEEVIAPIFTYVPLQLLAYHVALIKGTDVQDPR 600
Db 541 GGQLYVPADQAGFVSSDNMHHIEMPHVEEVIAPIFTYVPLQLLAYHVALIKGTDVQDPR 600
Qy 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609
RESULT 12
US-10-612-779-8
; Sequence 8, Application US/10612779
; Publication No. US20040091976A1
; GENERAL INFORMATION:
; APPLICANT: Deng, Ming-De
; APPLICANT: Angerer, J. David
; APPLICANT: Cyron, Don
; APPLICANT: Grund, Alan
; APPLICANT: Jerrell Jr., Thomas
; APPLICANT: Leanna, Candice
; APPLICANT: Mathre, Owen
; APPLICANT: Rosson, Reinhardt
; APPLICANT: Running, Jeff
; APPLICANT: Severson, Dave
; APPLICANT: Song, Linsheng
; APPLICANT: Massink, Sarah
; TITLE OF INVENTION: Process and Materials for Production of Glucosamine
; FILE REFERENCE: 3161-18-2
; CURRENT APPLICATION NUMBER: US/10/612,779
; CURRENT FILING DATE: 2003-07-01
; PRIOR APPLICATION NUMBER: 60/393,348
; PRIOR FILING DATE: 2002-07-01
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-612-779-8
Query Match 99.8%; Score 3079; DB 4; Length 609;
Best Local Similarity 99.8%; Pred. No. 1.8e-248;
Matches 608; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MCGIVGAIAQORDVAEILLEGRLRLLEYRGYDSAGLAVVDAGHMTLRRLRGKVQMLAQAAE 60
Db 1 MCGIVGAIAQORDVAEILLEGRLRLLEYRGYDSAGLAVVDAGHMTLRRLRGKVQMLAQAAE 60
Qy 61 EPHLHGCTGIAHTRWATHGEPSEVNAHPHVSEHIVVHNGIENHPELREBELKARGYTFV 120
Db 61 EPHLHGCTGIAHTRWATHGEPSEVNAHPHVSEHIVVHNGIENHPELREBELKARGYTFV 120
Qy 121 SETDTEVIAHLVNWELKQGGLTREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPLV 180
Db 121 SETDTEVIAHLVNWELKQGGLTREAVLRAIPQLRGAYGTVMDSRHPDITLLAARSGSPLV 180
Qy 181 IGLGMGNFTASDQALLPVTRRFIFLEEGDIAETRRSVNIFDKTGAEVKRODIESNLQ 240
Db 181 IGLGMGNFTASDQALLPVTRRFIFLEEGDIAETRRSVNIFDKTGAEVKRODIESNLQ 240
Qy 241 YDAGDKGIYRHYMQKEIYEOPNAIKNTLTGRISHGQVDLSELGPNADLLSKVEHIQILA 300
Db 241 YDAGDKGIYRHYMQKEIYEOPNAIKNTLTGRISHGQVDLSELGPNADLLSKVEHIQILA 300
Qy 301 CGTSYNSGMVSRVWFESLAGIPCDVEIASFPRYKSAVRNSLMITLQSGETADTLAGL 360

Db 301 COTSYNGMVSRVWFESLAGIPCDVEIASFEPRYKSAVRNLSMLITLSQSGETADTLA 360
QY 361 RLSKELGYLGSIAICNVPGSSIVRESDLALMTNAGTEIGVASTKAFQTOLTQTLVLLMLVAKL 420
Db 361 RLSKELGYLGSIAICNVPGSSIVRESDLALMTNAGTEIGVASTKAFQTOLTQTLVLLMLVAKL 420
QY 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMLSDQKRIEALAEFSDKHHLFLGRGDQYPIA 480
Db 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMLSDQKRIEALAEFSDKHHLFLGRGDQYPIA 480
QY 481 LEGALKLKEISYIHAEYAAGELKHGPLALIDAMPVIVVAPNNLEKLSNIEEVRAR 540
Db 481 LEGALKLKEISYIHAEYAAGELKHGPLALIDAMPVIVVAPNNLEKLSNIEEVRAR 540
QY 541 GGQLYVPADQAGFVSSDNMHIIEMPHVEEVIAPFYTPVPLQLLAYHVALIKGTVDQDPR 600
Db 541 GGQLYVPADQAGFVSSDNMHIIEMPHVEEVIAPFYTPVPLQLLAYHVALIKGTVDQDPR 600
QY 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 13
US-10-024-460-19
; Sequence 19, Application US/10024460
; Publication No. US20030044939A1
; GENERAL INFORMATION:
; APPLICANT: Berry, Alan
; APPLICANT: Burlingame, Richard P.
; APPLICANT: Millie, James R.
; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE
; FILE REFERENCE: 3161-18-C1
; CURRENT APPLICATION NUMBER: US/10/024,460
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/115,475
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/035,494
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-01-14
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-024-460-19

Query Match 99.5%; Score 3071; DB 4; Length 609;
Best Local Similarity 99.5%; Pred. No. 8.5e-248;
Matches 606; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MCGTGAIAQRDVAEILLEGRLRLEVRGYDSAGLAVVDAEGHMTLRLRLGKQVQLAAQAE 60
Db 1 MCGTGAIAQRDVAEILLEGRLRLEVRGYDSAGLAVVDAEGHMTLRLRLGKQVQLAAQAE 60

QY 61 EPHLHGCTGTAHRTWATHGEPSEVNAPHVSEHVVVHNGIIEHNEPLREELKARGYTFV 120
Db 61 EPHLHGCTGTAHRTWATHGEPSEVNAPHVSEHVVVHNGIIEHNEPLREELKARGYTFV 120

QY 121 SETDTEVIAHLVNWELKQGGTLEAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPVLV 180
Db 121 SETDTEVIAHLVNWELKQGGTLEAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPVLV 180

QY 181 IGLGMENFTASDQLALLPVTFRFIFLEEGDIAETRRSVNIFDKTGAEVKRODIESNLQ 240
Db 181 IGLGMENFTASDQLALLPVTFRFIFLEEGDIAETRRSVNIFDKTGAEVKRODIESNLQ 240

QY 241 YDAGDKGIVRHYMQKEIYQOPNAIKNTLTGRTSHGQVDLSLGNADLLSKVEHIQILA 300
Db 241 YDAGDKGIVRHYMQKEIYQOPNAIKNTLTGRTSHGQVDLSLGNADLLSKVEHIQILA 300

QY 301 COTSYNGMVSRVWFESLAGIPCDVEIASFEPRYKSAVRNLSMLITLSQSGETADTLA 360
Db 361 RLSKELGYLGSIAICNVPGSSIVRESDLALMTNAGTEIGVASTKAFQTOLTQTLVLLMLVAKL 420
Db 361 RLSKELGYLGSIAICNVPGSSIVRESDLALMTNAGTEIGVASTKAFQTOLTQTLVLLMLVAKL 420
QY 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMLSDQKRIEALAEFSDKHHLFLGRGDQYPIA 480
Db 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMLSDQKRIEALAEFSDKHHLFLGRGDQYPIA 480
QY 481 LEGALKLKEISYIHAEYAAGELKHGPLALIDAMPVIVVAPNNLEKLSNIEEVRAR 540
Db 481 LEGALKLKEISYIHAEYAAGELKHGPLALIDAMPVIVVAPNNLEKLSNIEEVRAR 540
QY 541 GGQLYVPADQAGFVSSDNMHIIEMPHVEEVIAPFYTPVPLQLLAYHVALIKGTVDQDPR 600
Db 541 GGQLYVPADQAGFVSSDNMHIIEMPHVEEVIAPFYTPVPLQLLAYHVALIKGTVDQDPR 600
QY 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

Db 301 COTSYNGMVSRVWFESLAGIPCDVEIASFEPRYKSAVRNLSMLITLSQSGETADTLA 360
QY 361 RLSKELGYLGSIAICNVPGSSIVRESDLALMTNAGTEIGVASTKAFQTOLTQTLVLLMLVAKL 420
Db 361 RLSKELGYLGSIAICNVPGSSIVRESDLALMTNAGTEIGVASTKAFQTOLTQTLVLLMLVAKL 420
QY 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMLSDQKRIEALAEFSDKHHLFLGRGDQYPIA 480
Db 421 SRLKGLDASIEHDIHVHGLQALPSRIEQMLSDQKRIEALAEFSDKHHLFLGRGDQYPIA 480
QY 481 LEGALKLKEISYIHAEYAAGELKHGPLALIDAMPVIVVAPNNLEKLSNIEEVRAR 540
Db 481 LEGALKLKEISYIHAEYAAGELKHGPLALIDAMPVIVVAPNNLEKLSNIEEVRAR 540
QY 541 GGQLYVPADQAGFVSSDNMHIIEMPHVEEVIAPFYTPVPLQLLAYHVALIKGTVDQDPR 600
Db 541 GGQLYVPADQAGFVSSDNMHIIEMPHVEEVIAPFYTPVPLQLLAYHVALIKGTVDQDPR 600
QY 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 14
US-10-612-779-4
; Sequence 4, Application US/10612779
; Publication No. US20040091976A1
; GENERAL INFORMATION:
; APPLICANT: Deng, Ming-De
; APPLICANT: Angerer, J. David
; APPLICANT: Cytron, Don
; APPLICANT: Grund, Alan
; APPLICANT: Jerrell Jr., Thomas
; APPLICANT: Leanna, Candice
; APPLICANT: Mathre, Owen
; APPLICANT: Rosson, Reinhardt
; APPLICANT: Running, Jeff
; APPLICANT: Severson, Dave
; APPLICANT: Song, Linsheng
; APPLICANT: Wassink, Sarah
; TITLE OF INVENTION: Process and Materials for Production of Glucosamine and N-Acetytle
; FILE REFERENCE: 3161-18-2
; CURRENT APPLICATION NUMBER: US/10/612,779
; CURRENT FILING DATE: 2003-07-01
; PRIOR APPLICATION NUMBER: 60/393,348
; PRIOR FILING DATE: 2002-07-01
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-612-779-4

Query Match 99.5%; Score 3071; DB 4; Length 609;
Best Local Similarity 99.5%; Pred. No. 8.5e-248;
Matches 606; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MCGTGAIAQRDVAEILLEGRLRLEVRGYDSAGLAVVDAEGHMTLRLRLGKQVQLAAQAE 60
Db 1 MCGTGAIAQRDVAEILLEGRLRLEVRGYDSAGLAVVDAEGHMTLRLRLGKQVQLAAQAE 60

QY 61 EPHLHGCTGTAHRTWATHGEPSEVNAPHVSEHVVVHNGIIEHNEPLREELKARGYTFV 120
Db 61 EPHLHGCTGTAHRTWATHGEPSEVNAPHVSEHVVVHNGIIEHNEPLREELKARGYTFV 120

QY 121 SETDTEVIAHLVNWELKQGGTLEAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPVLV 180
Db 121 SETDTEVIAHLVNWELKQGGTLEAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPVLV 180

QY 181 IGLGMENFTASDQLALLPVTFRFIFLEEGDIAETRRSVNIFDKTGAEVKRODIESNLQ 240
Db 181 IGLGMENFTASDQLALLPVTFRFIFLEEGDIAETRRSVNIFDKTGAEVKRODIESNLQ 240

```
Db      181  IGLGMGFNIASDQLALLPVTRRFIFLEEGDIAEITRRSVNIFDKTGAEVKQDIESNLQ 240
Qy      241  YDAGDKGIYRHYMQKEIYEOPNAIKNTLTGRISHGQVDLSLGNADLLEKSKVEHIQILA 300
Db      241  YDAGDKGIYRHYMQKEIYEOPNAIKNTLTGRISHGQVDLSLGNADLLEKSKVEHIQILA 300
Qy      301  CGTSYNSGMVSRVWFESLAGIPCDVEIASFPRYKSAVRNLSLMTLSQSGETADTLA 360
Db      301  CGTSYNSGMVSRVWFESLAGIPCDVEIASFPRYKSAVRNLSLMTLSQSGETADTLA 360
Qy      361  RLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFPTTQTLVLLMLVAKL 420
Db      361  RLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFPTTQTLVLLMLVAKL 420
Qy      421  SRLKGLDASTIEHDI VHGLQALPSRIEOMLSQDKRIEALAEFSDKHALFLRGDOVPIA 480
Db      421  SRLKGLDASTIEHDI VHGLQALPSRIEOMLSQDKRIEALAEFSDKHALFLRGDOVPIA 480
Qy      481  LEGALKLKEISYTHAEYAAAGELKHGFLALIDADMPVIVVAPNNELLEKLKSNIEEVRAR 540
Db      481  LEGALKLKEISYTHAEYAAAGELKHGFLALIDADMPVIVVAPNNELLEKLKSNIEEVRAR 540
Qy      541  GGOLYVPADQDAGFVSSDNHIIEMPHVEVIAPIFYTPVPLQLLAYHVALIKGTDVDDQR 600
Db      541  GGOLYVPADQDAGFVSSDNHIIEMPHVEVIAPIFYTPVPLQLLAYHVALIKGTDVDDQR 600
Qy      601  NLAQSVTVE 609
Db      601  NLAQSVTVE 609

RESULT 15
US-10-024-460-22
; Sequence 22, Application US/10024460
; Publication No. US20030044939A1
; GENERAL INFORMATION:
; APPLICANT: Berry, Alan
; APPLICANT: Mulline, James R.
; APPLICANT: Burlingame, Richard P.
; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE
; CURRENT APPLICATION NUMBER: US/10/024,460
; FILE REFERENCE: 3161-18-C1
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/115,475
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/035,494
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-01-14
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-024-460-22

Query Match      99.4%; Score 3068; DB 4; Length 609;
Best Local Similarity 99.5%; Pred. No. 1.5e-247;
Matches 606; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      1  MCGIVGAIQRDVAEILLEGRLREYRGYDSAGLAVVDAEGHMTLRRLRGKVQMLAQAAE 60
Db      1  MCGIVGAIQRDVAEILLEGRLREYRGYDSAGLAVVDTGEGHMTLRRLRGKVQMLAQAAE 60
Qy      61  EHPHGGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEPLREELKARGYTFV 120
Db      61  EHPHGGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIIEHNEPLREELKARGYTFV 120
Qy      121  SETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDTLAARSGSPLV 180
Db      121  SETDTEVIAHLVNWELKQGGTLREAVLRAIPQLRGAYGTVMDSRHPDTLAARSGSPLV 180
Qy      181  IGLGMGFNIASDQLALLPVTRRFIFLEEGDIAEITRRSVNIFDKTGAEVKQDIESNLQ 240
Db      181  IGLGMGFNIASDQLALLPVTRRFIFLEEGDIAEITRRSVNIFDKTGAEVKQDIESNLQ 240
```

Search completed: August 7, 2006, 09:28:24
Job time : 95 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 7, 2006, 09:27:53 ; Search time 156 Seconds
(without alignments)
261.222 Million cell updates/sec

Title: US-10-612-779-2

Perfect score: 3086

Sequence: 1 MCGIVGAIQRDVAEILLEG.....LIRGTDVDPNRLAKSVTVE 609

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 236815 seqs, 66914042 residues

Total number of hits satisfying chosen parameters: 236815

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA New:

- 1: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 2: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 3: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 4: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 5: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 6: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 7: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 8: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|---------------------------------------|
| 1 | 3086 | 100.0 | 609 | 7 | US-11-245-473-16 Sequence 16, Appl |
| 2 | 3080 | 99.8 | 609 | 7 | US-11-245-473-28 Sequence 28, Appl |
| 3 | 3080 | 99.8 | 609 | 7 | US-11-245-473-31 Sequence 31, Appl |
| 4 | 3079 | 99.8 | 609 | 7 | US-11-245-473-25 Sequence 25, Appl |
| 5 | 3071 | 99.5 | 609 | 7 | US-11-245-473-19 Sequence 19, Appl |
| 6 | 3068 | 99.4 | 609 | 7 | US-11-245-473-22 Sequence 22, Appl |
| 7 | 1150 | 37.3 | 601 | 6 | US-10-471-571A-2268 Sequence 2268, Ap |
| 8 | 1109 | 35.9 | 681 | 6 | US-10-480-021-6 Sequence 6, Appli |
| 9 | 1109 | 35.9 | 681 | 6 | US-10-480-021-7 Sequence 7, Appli |
| 10 | 1103.5 | 35.8 | 682 | 6 | US-10-480-021-8 Sequence 8, Appli |
| 11 | 1089.5 | 35.3 | 680 | 6 | US-10-953-349-34174 Sequence 34174, A |
| 12 | 1078 | 34.9 | 697 | 6 | US-10-449-902-38427 Sequence 38427, A |
| 13 | 808 | 26.2 | 491 | 6 | US-10-953-349-34175 Sequence 34175, A |
| 14 | 807 | 26.2 | 488 | 6 | US-10-953-349-34176 Sequence 34176, A |
| 15 | 589.5 | 19.1 | 464 | 6 | US-10-488-015-17 Sequence 17, Appl |
| 16 | 557.5 | 18.1 | 286 | 6 | US-10-449-902-36965 Sequence 36965, A |
| 17 | 252 | 8.2 | 541 | 6 | US-10-449-902-54059 Sequence 54059, A |
| 18 | 240 | 7.8 | 501 | 6 | US-10-471-571A-5078 Sequence 5078, Ap |
| 19 | 222.5 | 7.2 | 563 | 7 | US-11-330-403-6080 Sequence 6080, Ap |
| 20 | 188.5 | 6.1 | 600 | 6 | US-10-449-902-38704 Sequence 38704, A |
| 21 | 186 | 6.0 | 541 | 7 | US-11-330-403-13759 Sequence 13759, A |
| 22 | 186 | 6.0 | 544 | 7 | US-11-330-403-6869 Sequence 6869, Ap |
| 23 | 184.5 | 6.0 | 78 | 6 | US-10-488-015-16 Sequence 16, Appl |
| 24 | 183.5 | 5.9 | 610 | 7 | US-11-330-403-1786 Sequence 1786, Ap |
| 25 | 183 | 5.9 | 525 | 7 | US-11-330-403-17502 Sequence 17502, A |

ALIGNMENTS

RESULT 1

US-11-245-473-16

; Sequence 16, Application US/11245473

; Publication No. US20060094085A1

; GENERAL INFORMATION:

; APPLICANT: Berry, Alan

; APPLICANT: Burlingame, Richard P.

; APPLICANT: Mullis, James R.

; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE

; FILE REFERENCE: 3161-18-C1

; CURRENT APPLICATION NUMBER: US/11/245,473

; CURRENT FILING DATE: 2005-10-05

; PRIOR APPLICATION NUMBER: US/10/024,460

; PRIOR FILING DATE: 2001-12-17

; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/115,475

; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-15

; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/035,494

; PRIOR FILING DATE: EARLIER FILING DATE: 1997-01-14

; NUMBER OF SEQ ID NOS: 31

; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 16

; LENGTH: 609

; TYPE: PRT

; ORGANISM: Escherichia coli

US-11-245-473-16

Query Match 100.0%; Score 3086; DB 7; Length 609;

Best Local Similarity 100.0%; Pred. No. 9.8e-221;

Matches 609; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MCGIVGAIQRDVAEILLEGRLRLEVRGYSAGLVVDAEGHMTLRRLRKGVMLOAAAE 60

Db 1 MCGIVGAIQRDVAEILLEGRLRLEVRGYSAGLVVDAEGHMTLRRLRKGVMLOAAAE 60

Qy 61 EPHLGGTGTIAHTRWATHGSEVNAPHPVSEHVVVHNGIIEHHEPLREELKARGYTFV 120

Db 61 EPHLGGTGTIAHTRWATHGSEVNAPHPVSEHVVVHNGIIEHHEPLREELKARGYTFV 120

Qy 121 SETDTEVIAHLVNWELKQGGTLREAVLRALPOLRGAYGTVMDSRHPDILLARSGSLV 180

Db 121 SETDTEVIAHLVNWELKQGGTLREAVLRALPOLRGAYGTVMDSRHPDILLARSGSLV 180

Qy 181 IGLGMENFTASDQLALLPVTRRFIFLEEGDIAEITRRSVNIPDKTGAEVKRODIESNLQ 240

Db 181 IGLGMENFTASDQLALLPVTRRFIFLEEGDIAEITRRSVNIPDKTGAEVKRODIESNLQ 240

Qy 241 YDAGDKGIYRHYMKEIYEQPNAIKNTLTGRISHGQVDLSLGNADLLSKVEHIQILA 300

Db 241 YDAGDKGIYRHYMKEIYEQPNAIKNTLTGRISHGQVDLSLGNADLLSKVEHIQILA 300

Db 241 YDAGDKGIYRHYMQKEIYEQPNAIKNTLTGRISHGQVDLSLGPNADELLSKVEHIQILA 300
Qy 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASSEFRYKSAVRNSMLITLSQSGETADTLAGL 360
Db 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASSEFRYKSAVRNSMLITLSQSGETADTLAGL 360
Qy 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFITQTLTVLLMLVAKL 420
Db 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFITQTLTVLLMLVAKL 420
Qy 421 SRLKGLDASIEHDIYHGLQALPSRIEQMLSQDKRIEALAEFSDKHALFLGRGDQYPIA 480
Db 421 SRLKGLDASIEHDIYHGLQALPSRIEQMLSQDKRIEALAEFSDKHALFLGRGDQYPIA 480
Qy 481 LEGALKLKEISYIHAEYAAGELKHGPLALIDAMPVIVVAPNNELLEKLKSNIEVRAR 540
Db 481 LEGALKLKEISYIHAEYAAGELKHGPLALIDAMPVIVVAPNNELLEKLKSNIEVRAR 540
Qy 541 GGQLYVFADQDAGFVSSDNMHIIEMPHVEEVIAPFIVTVPLQLLAYHVALIKGTDVDDPR 600
Db 541 GGQLYVFADQDAGFVSSDNMHIIEMPHVEEVIAPFIVTVPLQLLAYHVALIKGTDVDDPR 600
Qy 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 2

US-11-245-473-28
; Sequence 28, Application US/11245473
; Publication No. US20060094085A1
; GENERAL INFORMATION:
; APPLICANT: Berry, Alan
; APPLICANT: Burlingame, Richard P.
; APPLICANT: Millis, James R.
; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE
; FILE REFERENCE: 3161-18-C1
; CURRENT APPLICATION NUMBER: US/11/245,473
; PRIOR FILING DATE: 2005-10-05
; PRIOR APPLICATION NUMBER: US/10/024,460
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/115,475
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/035,494
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-01-14
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 28
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
US-11-245-473-28

Query Match 99.8%; Score 3080; DB 7; Length 609;
Best Local Similarity 99.8%; Pred. No. 2.7e-220;
Matches 608; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MCGIVGAIAQRDVAEILLEGRLRLEYRGYDSAGLAVVDAEGHMTLRRLRGKVMQLAQAAE 60
Db 1 MCGIVGAIAQRDVAEILLEGRLRLEYRGYDSAGLAVVDAEGHMTLRRLRGKVMQLAQAAE 60
Qy 61 EHLPGGTGTAHTRWATHGEPSEVNAPHVSEHIVVHNGIIEENHEPLREELKARGYTFV 120
Db 61 EHLPGGTGTAHTRWATHGEPSEVNAPHVSEHIVVHNGIIEENHEPLREELKARGYTFV 120
Qy 121 SETDTEVIAHLVNWELKQGTTREAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPLV 180
Db 121 SETDTEVIAHLVNWELKQGTTREAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPLV 180
Qy 181 IGLGMGENFIASDQALLPVTRRFIFLEEGDIAETRRSVNIFDKTGAEVKQDIESNLQ 240
Db 181 IGLGMGENFIASDQALLPVTRRFIFLEEGDIAETRRSVNIFDKTGAEVKQDIESNLQ 240

Qy 241 YDAGDKGIYRHYMQKEIYEQPNAIKNTLTGRISHGQVDLSLGPNADELLSKVEHIQILA 300
Db 241 YDAGDKGIYRHYMQKEIYEQPNAIKNTLTGRISHGQVDLSLGPNADELLSKVEHIQILA 300
Qy 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASSEFRYKSAVRNSMLITLSQSGETADTLAGL 360
Db 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASSEFRYKSAVRNSMLITLSQSGETADTLAGL 360
Qy 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFITQTLTVLLMLVAKL 420
Db 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFITQTLTVLLMLVAKL 420
Qy 421 SRLKGLDASIEHDIYHGLQALPSRIEQMLSQDKRIEALAEFSDKHALFLGRGDQYPIA 480
Db 421 SRLKGLDASIEHDIYHGLQALPSRIEQMLSQDKRIEALAEFSDKHALFLGRGDQYPIA 480
Qy 481 LEGALKLKEISYIHAEYAAGELKHGPLALIDAMPVIVVAPNNELLEKLKSNIEVRAR 540
Db 481 LEGALKLKEISYIHAEYAAGELKHGPLALIDAMPVIVVAPNNELLEKLKSNIEVRAR 540
Qy 541 GGQLYVFADQDAGFVSSDNMHIIEMPHVEEVIAPFIVTVPLQLLAYHVALIKGTDVDDPR 600
Db 541 GGQLYVFADQDAGFVSSDNMHIIEMPHVEEVIAPFIVTVPLQLLAYHVALIKGTDVDDPR 600
Qy 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 3

US-11-245-473-31
; Sequence 31, Application US/11245473
; Publication No. US20060094085A1
; GENERAL INFORMATION:
; APPLICANT: Berry, Alan
; APPLICANT: Burlingame, Richard P.
; APPLICANT: Millis, James R.
; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE
; FILE REFERENCE: 3161-18-C1
; CURRENT APPLICATION NUMBER: US/11/245,473
; PRIOR FILING DATE: 2005-10-05
; PRIOR APPLICATION NUMBER: US/10/024,460
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/115,475
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/035,494
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-01-14
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 31
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
US-11-245-473-31

Query Match 99.8%; Score 3080; DB 7; Length 609;
Best Local Similarity 99.8%; Pred. No. 2.7e-220;
Matches 608; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MCGIVGAIAQRDVAEILLEGRLRLEYRGYDSAGLAVVDAEGHMTLRRLRGKVMQLAQAAE 60
Db 1 MCGIVGAIAQRDVAEILLEGRLRLEYRGYDSAGLAVVDAEGHMTLRRLRGKVMQLAQAAE 60
Qy 61 EHLPGGTGTAHTRWATHGEPSEVNAPHVSEHIVVHNGIIEENHEPLREELKARGYTFV 120
Db 61 EHLPGGTGTAHTRWATHGEPSEVNAPHVSEHIVVHNGIIEENHEPLREELKARGYTFV 120
Qy 121 SETDTEVIAHLVNWELKQGTTREAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPLV 180
Db 121 SETDTEVIAHLVNWELKQGTTREAVLRAIPQLRGAYGTVMDSRHPDTLLAARSGSPLV 180
Qy 181 IGLGMGENFIASDQALLPVTRRFIFLEEGDIAETRRSVNIFDKTGAEVKQDIESNLQ 240
Db 181 IGLGMGENFIASDQALLPVTRRFIFLEEGDIAETRRSVNIFDKTGAEVKQDIESNLQ 240

Db 181 IGLGMGNFIASDQALLPVTRRFIFLEEGDIABITRRSVNIPDKTGAEVKRODIESNLQ 240
Qy 241 YDAGDKGIYRHYMQKEIYEQNAIKNTLTGRISHGQVDLSLGNADLLSKVHIQILA 300
Db 241 YDAGDKGIYRHYMQKEIYEQNAIKNTLTGRISHGQVDLSLGNADLLSKVHIQILA 300
Qy 301 CGTSYNGMSRYWFESLAGIPCDVEIASFPRKSAVRRNSLMTLTSQSGETADTLAGL 360
Db 301 CGTSYNGMSRYWFESLAGIPCDVEIASFPRKSAVRRNSLMTLTSQSGETADTLAGL 360
Qy 361 RLSKELGYLSLAI CNVPGSSLVRESLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 420
Db 361 RLSKELGYLSLAI CNVPGSSLVRESLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 420
Qy 421 SRLKGLDASIEHDIVHGLQALPSRIEQMLSDOKRIEALAEDFSOKHHALFLGRGDQYPIA 480
Db 421 SRLKGLDASIEHDIVHGLQALPSRIEQMLSDOKRIEALAEDFSOKHHALFLGRGDQYPIA 480
Qy 481 LEGALKLKEISYTHAEAYAAGELKHGPLALIDAMPVIVVAPNNLEKLSNIEEVRRAR 540
Db 481 LEGALKLKEISYTHAEAYAAGELKHGPLALIDAMPVIVVAPNNLEKLSNIEEVRRAR 540
Qy 541 GGQLYVFADQDAGFVSSDNMHIIEMPHVEEVIAPIFTVPQLLAYHVALIKGTDVQDPR 600
Db 541 GGQLYVFADQDAGFVSSDNMHIIEMPHVEEVIAPIFTVPQLLAYHVALIKGTDVQDPR 600
Qy 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 4

US-11-245-473-25
; Sequence 25, Application US/11245473
; Publication No. US20060094085A1
; GENERAL INFORMATION:
; APPLICANT: Berry, Alan
; APPLICANT: Burlingame, Richard P.
; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE
; FILE REFERENCE: 3161-18-C1
; CURRENT APPLICATION NUMBER: US/11/245,473
; PRIOR FILING DATE: 2005-10-05
; PRIOR FILING DATE: 2001-12-17
; PRIOR FILING DATE: 2001-12-17
; PRIOR FILING DATE: EARLIER APPLICATION NUMBER: 09/115,475
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-15
; PRIOR FILING DATE: EARLIER APPLICATION NUMBER: 60/035,494
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
US-11-245-473-25

Query Match 99.8%; Score 3079; DB 7; Length 609;
Best Local Similarity 99.8%; Pred. No. 3.2e-220;
Matches 608; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MCGIVGAIAORDVAEILLEGRLRLEYRGYDSAGLAVVDAEGHMTLRLRLGKVQMLAQAAE 60
Db 1 MCGIVGAIAORDVAEILLEGRLRLEYRGYDSAGLAVVDAEGHMTLRLRLGKVQMLAQAAE 60
Qy 61 EPHLHGCTGTAHTRWATHGEPSEVNAHPHYSEHIVVHNGIENHEPRLBELKARGYTFV 120
Db 61 EPHLHGCTGTAHTRWATHGEPSEVNAHPHYSEHIVVHNGIENHEPRLBELKARGYTFV 120
Qy 121 SETDTEVIAHLVNWELKOGGTLREAVLRAIPQLRGAYGTVMDSRHPDPTLLAARSGSPLV 180
Db 121 SETDTEVIAHLVNWELKOGGTLREAVLRAIPQLRGAYGTVMDSRHPDPTLLAARSGSPLV 180

Qy 181 IGLGMGNFIASDQALLPVTRRFIFLEEGDIABITRRSVNIPDKTGAEVKRODIESNLQ 240
Db 181 IGLGMGNFIASDQALLPVTRRFIFLEEGDIABITRRSVNIPDKTGAEVKRODIESNLQ 240
Qy 241 YDAGDKGIYRHYMQKEIYEQNAIKNTLTGRISHGQVDLSLGNADLLSKVHIQILA 300
Db 241 YDAGDKGIYRHYMQKEIYEQNAIKNTLTGRISHGQVDLSLGNADLLSKVHIQILA 300
Qy 301 CGTSYNGMSRYWFESLAGIPCDVEIASFPRKSAVRRNSLMTLTSQSGETADTLAGL 360
Db 301 CGTSYNGMSRYWFESLAGIPCDVEIASFPRKSAVRRNSLMTLTSQSGETADTLAGL 360
Qy 361 RLSKELGYLSLAI CNVPGSSLVRESLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 420
Db 361 RLSKELGYLSLAI CNVPGSSLVRESLALMTNAGTEIGVASTKFTTQLTVLLMLVAKL 420
Qy 421 SRLKGLDASIEHDIVHGLQALPSRIEQMLSDOKRIEALAEDFSOKHHALFLGRGDQYPIA 480
Db 421 SRLKGLDASIEHDIVHGLQALPSRIEQMLSDOKRIEALAEDFSOKHHALFLGRGDQYPIA 480
Qy 481 LEGALKLKEISYTHAEAYAAGELKHGPLALIDAMPVIVVAPNNLEKLSNIEEVRRAR 540
Db 481 LEGALKLKEISYTHAEAYAAGELKHGPLALIDAMPVIVVAPNNLEKLSNIEEVRRAR 540
Qy 541 GGQLYVFADQDAGFVSSDNMHIIEMPHVEEVIAPIFTVPQLLAYHVALIKGTDVQDPR 600
Db 541 GGQLYVFADQDAGFVSSDNMHIIEMPHVEEVIAPIFTVPQLLAYHVALIKGTDVQDPR 600
Qy 601 NLAKSVTVE 609
Db 601 NLAKSVTVE 609

RESULT 5

US-11-245-473-19
; Sequence 19, Application US/11245473
; Publication No. US20060094085A1
; GENERAL INFORMATION:
; APPLICANT: Berry, Alan
; APPLICANT: Burlingame, Richard P.
; APPLICANT: Millis, James R.
; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE
; FILE REFERENCE: 3161-18-C1
; CURRENT APPLICATION NUMBER: US/11/245,473
; PRIOR FILING DATE: 2005-10-05
; PRIOR FILING DATE: 2001-12-17
; PRIOR FILING DATE: 2001-12-17
; PRIOR FILING DATE: EARLIER APPLICATION NUMBER: 09/115,475
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-15
; PRIOR FILING DATE: EARLIER APPLICATION NUMBER: 60/035,494
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
US-11-245-473-19

Query Match 99.5%; Score 3071; DB 7; Length 609;
Best Local Similarity 99.5%; Pred. No. 1.3e-219;
Matches 606; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 MCGIVGAIAORDVAEILLEGRLRLEYRGYDSAGLAVVDAEGHMTLRLRLGKVQMLAQAAE 60
Db 1 MCGIVGAIAORDVAEILLEGRLRLEYRGYDSAGLAVVDAEGHMTLRLRLGKVQMLAQAAE 60
Qy 61 EPHLHGCTGTAHTRWATHGEPSEVNAHPHYSEHIVVHNGIENHEPRLBELKARGYTFV 120
Db 61 EPHLHGCTGTAHTRWATHGEPSEVNAHPHYSEHIVVHNGIENHEPRLBELKARGYTFV 120
Qy 121 SETDTEVIAHLVNWELKOGGTLREAVLRAIPQLRGAYGTVMDSRHPDPTLLAARSGSPLV 180

Db 121 SETDTEVIAHLVNMWELKQGGTLREAVLRAIPQLRGAYGTVIMDSRHPDPTLLAARSGSPLV 180
Qy 181 IGLGNGENFIASDQALALLPVTRRFIFLEEGDIAETIRRSVNIIPKTAQAEVKRODIESNLQ 240
Db 181 IGLGNGENFIASDQALALLPVTRRFIFLEEGDIAETIRRSVNIIPKTAQAEVKRODIESNLQ 240
Qy 241 YDAGDKGIYRHYMQKEIYEQPNAINKNTLTGRISHGQVDLSLSELGNADLLSKVEHIQILA 300
Db 241 YDAGDKGIYRHYMQKEIYEQPNAINKNTLTGRISHGQVDLSLSELGNADLLSKVEHIQILA 300
Qy 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASFEPYRKSAVRRNSLMTITLSQSGETADTLA 360
Db 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASFEPYRKSAVRRNSLMTITLSQSGETADTLA 360
Qy 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFITQTLTVLLMLVAKL 420
Db 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFITQTLTVLLMLVAKL 420
Qy 421 SRLKGLDASIEHDIIVHGLQALPSRIEQMLPDKRIEALAEFSDKHALFLGRGDQYPIA 480
Db 421 SRLKGLDASIEHDIIVHGLQALPSRIEQMLPDKRIEALAEFSDKHALFLGRGDQYPIA 480
Qy 481 LEGALKLKEISYIHAAYAGELKHGPIALIDAMPVIVVAPNNELLEKLKSNIEEVRAR 540
Db 481 LEGALKLKEISYIHAAYAGELKHGPIALIDAMPVIVVAPNNELLEKLKSNIEEVRAR 540
Qy 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPFIVTVPLQLLAYHVALIKGTDVQDR 600
Db 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPFIVTVPLQLLAYHVALIKGTDVQDR 600
Qy 601 NLAQSVTVE 609
Db 601 NLAQSVTVE 609

RESULT 6

US-11-245-473-22
; Sequence 22, Application US/11245473
; Publication No. US20060094085A1
; GENERAL INFORMATION:
; APPLICANT: Berry, Alan
; APPLICANT: Burlingame, Richard P.
; APPLICANT: Millis, James R.
; TITLE OF INVENTION: PROCESS AND MATERIALS FOR PRODUCTION OF GLUCOSAMINE
; FILE REFERENCE: 3161-18-C1
; CURRENT APPLICATION NUMBER: US/11/245,473
; CURRENT FILING DATE: 2005-10-05
; PRIOR FILING DATE: 2001-12-17
; PRIOR FILING DATE: 1997-01-14
; PRIOR FILING DATE: 1998-07-15
; PRIOR FILING DATE: 1997-01-14
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Escherichia coli
US-11-245-473-22

Query Match 99.4%; Score 3068; DB 7; Length 609;
Best Local Similarity 99.5%; Pred. No. 2,1e-219;
Matches 606; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MCGIVGAIAQRDVAEILLGLRLLEYRGYDSAGLAVVDAGHMTLRRLRGKVMQMLAAQAE 60
Db 1 MCGIVGAIAQRDVAEILLGLRLLEYRGYDSAGLAVVDAGHMTLRRLRGKVMQMLAAQAE 60
Qy 61 EHPHGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIENHEPRLREELKARGYTFV 120
Db 61 EHPHGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIENHEPRLREELKARGYTFV 120

Qy 121 SETDTEVIAHLVNMWELKQGGTLREAVLRAIPQLRGAYGTVIMDSRHPDPTLLAARSGSPLV 180
Db 121 SETDTEVIAHLVNMWELKQGGTLREAVLRAIPQLRGAYGTVIMDSRHPDPTLLAARSGSPLV 180
Qy 181 IGLGNGENFIASDQALALLPVTRRFIFLEEGDIAETIRRSVNIIPKTAQAEVKRODIESNLQ 240
Db 181 IGLGNGENFIASDQALALLPVTRRFIFLEEGDIAETIRRSVNIIPKTAQAEVKRODIESNLQ 240
Qy 241 YDAGDKGIYRHYMQKEIYEQPNAINKNTLTGRISHGQVDLSLSELGNADLLSKVEHIQILA 300
Db 241 YDAGDKGIYRHYMQKEIYEQPNAINKNTLTGRISHGQVDLSLSELGNADLLSKVEHIQILA 300
Qy 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASFEPYRKSAVRRNSLMTITLSQSGETADTLA 360
Db 301 CGTSYNSGMVSRYPFESLAGIPCDVEIASFEPYRKSAVRRNSLMTITLSQSGETADTLA 360
Qy 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFITQTLTVLLMLVAKL 420
Db 361 RLSKELGYLGSIAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFITQTLTVLLMLVAKL 420
Qy 421 SRLKGLDASIEHDIIVHGLQALPSRIEQMLPDKRIEALAEFSDKHALFLGRGDQYPIA 480
Db 421 SRLKGLDASIEHDIIVHGLQALPSRIEQMLPDKRIEALAEFSDKHALFLGRGDQYPIA 480
Qy 481 LEGALKLKEISYIHAAYAGELKHGPIALIDAMPVIVVAPNNELLEKLKSNIEEVRAR 540
Db 481 LEGALKLKEISYIHAAYAGELKHGPIALIDAMPVIVVAPNNELLEKLKSNIEEVRAR 540
Qy 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPFIVTVPLQLLAYHVALIKGTDVQDR 600
Db 541 GGQLYVFADQDAGFVSSDNMHIEMPHVEEVIAPFIVTVPLQLLAYHVALIKGTDVQDR 600
Qy 601 NLAQSVTVE 609
Db 601 NLAQSVTVE 609

RESULT 7

US-10-471-571A-2268
; Sequence 2268, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927WO
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 2268
; LENGTH: 601
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-471-571A-2268

Query Match 37.3%; Score 1150; DB 6; Length 601;
Best Local Similarity 40.8%; Pred. No. 3.2e-77;
Matches 252; Conservative 126; Mismatches 214; Indels 26; Gaps 10;

Qy 1 MCGIVGAIAQRDVAEILLGLRLLEYRGYDSAGLAVVDAGHMTLRRLRGKVMQMLAAQAE 60
Db 1 MCGIVGAIAQRDVAEILLGLRLLEYRGYDSAGLAVVDAGHMTLRRLRGKVMQMLAAQAE 59
Qy 61 EHPHGGTGIAHTRWATHGEPSEVNAHPVSEHIVVHNGIENHEPRLREELKARGYTFV 118
Db 60 SSDPDGPGVIGCHTRWATHGVNHNHNSHPQSSNGRFTLVHNGVIENYEELKGEY-LQGV 118
Qy 119 FVSTDTTEVIAHLVNMWELKQGGTLREAVLRAIPQLRGAYGTVIMDSRHPDPTLLAARSGSP 178
Db 119 FISTDTTEVIVQLVEYPSNQGLSTEEFTKVSLLHGSYALGLDLDAEDKOTIYVAKNKP 178

| | | | |
|----|-----|--|-----|
| Qy | 179 | LVIGLHGENFIASQJALLPVTRRFFLEBGDTAETRRSVNFDTKTAEGAEVKQODIESN | 238 |
| Db | 179 | LLGVGEGFNVIASDALAMLQVTSYKEIHDEHIVIVKKOEVIIKDADGNVVERDSYIAE | 238 |
| Qy | 239 | LOYDAGDKGIYRHYMQKEIYEQPNAIKNTL-----TGRISHGQVDLSGELGNADDELLSK | 292 |
| Db | 239 | IDASDAEKGVVYAHYMLKEIHQPAVMRRIIQEQDAEGNLKIDQDIIND-----VKE | 290 |
| Qy | 293 | VEHIQIIACGTSYNSGMVSRVYWFESLAGICPQVEIAESFPRYKSAVRNSLMITLSQSGE | 352 |
| Db | 291 | ADRIYVIAAGTSYHAGLVKGFEKFKWAGVPTVEHVASEFVYVNMPLSEKPLFVYISQSGE | 350 |
| Qy | 353 | TADTLAGLRLSKELGYLGSLAI CNVPGSSLVRESDLAMTNAGTEIGVASTKAPTQLTV | 412 |
| Db | 351 | TADSRVAVLVTNKLCH-KSLTITVNVAGSTLSREADHTLLHAGFEIAVASTKAYTAQTAIV | 409 |
| Qy | 413 | LLMLVAKLSRLKGLDASIEHDIHVHGLQALPSRIEQMLSQDKRIIFALAEFSD-KHHAUFL | 471 |
| Db | 410 | LSILSQIVAAHEHGREADI--DLLRELAKVTTAIEAIVDDAPIMEQIATDFLETTNRNAPFI | 467 |
| Qy | 472 | GRGDOYIALLEGALKLKEISYIHAEEVAAGELKHGPLALIDADMPVIVVAPNNELLEKL | 531 |
| Db | 468 | GRITIDYNSLEGALKLKEISYIQAGFAGGELKHGTIALIEEGTPVVGATQOEKNLSIR | 527 |
| Qy | 532 | SNIEEVARGQGLVYFADQDQAGFVSSDNMMHIIEMPHVEEVTAPIFYTPLOLLAYHVALI | 591 |
| Db | 528 | GNVKEVVARGAHPCCIISME--GLEKEGDTYVI--PHVHELLTPLVSVVALQLISYYAALH | 583 |
| Qy | 592 | KGTDVQDPRNLAKSVTVE | 609 |
| Db | 584 | RDLVDKPRNLAKSVTVE | 601 |

```

RESULT 8
US-10-480-021-6
; Sequence 6, Application US/10480021
; Publication No. US20060121457A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: GFATS AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-068
; CURRENT APPLICATION NUMBER: US/10/480,021
; CURRENT FILING DATE: 2003-12-04
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 681
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-480-021-6

```

[illegible]

```

RESULT 9
US-10-480-021-7
; Sequence 7, Application US/10480021
; Publication No. US20060121457A1
; GENERAL INFORMATION:
; APPLICANT: EXELUXIS, INC.
; TITLE OF INVENTION: GFATs AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-068
; CURRENT APPLICATION NUMBER: US/10/480,021
; CURRENT FILING DATE: 2003-12-04
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 681
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-480-021-7

```

```
Qy 95 VVHNGIENHEPLREELKARGYTFVSETDTEVIAHLVNM-----ELKQGGTLREAVLRAI 150
Db 119 IVIHNGIITNYKDLKFKLESKYDFESETDTEIAKLKVMYDNRESQDTSFTTLVERVI 178
Qy 151 POLRGAYGTVMDSRHPDPTLLAARSGSPVLIGLGMG-----
Db 179 QOLEGAFALVPKSVHFFGQAVGTRRGSPLLIGVRSKHLSTDHPIILYRTGDKKGSCNL 238
Qy 187 -----ENFIASQOLALLPVRTRFFIEEGDIAEIT--RRSVNIIPDKTG 227
Db 239 SRVDSITCLPVEKAVEYFASDASAVIEHTNRVIFLEDDDDAAVVDGRLSHRIKRTA 298
Qy 228 AEVRKQDIESNLOVDAGD--KGIYRHYMQKEIYEQPNAIKNTLTGRISHQVDSLGPKN 285
Db 299 GDHPGRAVQT-LQMELOQIMKGNFSSFMQKEIFEQPESVVNTWRGVNPDYTVN-LGGL 356
Qy 286 ADEL--LSKVEHIQILACGTSYNSGMVSRVWFSLAGIPCDVIEASEFRYKSAVRNSL 343
Db 357 KDHIKEIQRCRRLLIACGTSYHAGVATROVLEELTELPVMVELASDFLDNRNTPVFRDDV 416
Qy 344 MITLSQSGEATDTLAGRLSKELGYLGLSLAICNVPGSSLVRESDLALMTNAGTEIGVAST 403
Db 417 CFFLSQSGEATDTLMGLRYCKRGAL--TVGTTNTVSGSISRETDGCVHINAGPEIGVAST 475
Qy 404 KAFPTQLTVLLMLVAKLSRLKGLDASIEHDIVHGLQALPSRIEQMLSQDKRIEALAEDFS 463
Db 476 KAYTSQFVSLVMPALWMCDDRISSQOERRKEIMLGLKELPLDILKEVLSMDDEIQKLATELY 535
Qy 464 DKHIALPLRGDDQVPIALEGALKLKEISYTHASAYAGELKHGPLALIDADMVPIVAPN 523
Db 536 HQKSVLMGRGYHATCLEGALKIKEITYMHSEGILAGELKHGPLALVDKLMPEVIMIIR 595
Qy 524 NELLEKLKSNEEVRARGGQLYVPADQDAGFVSSDNHIIEMPHVEEVIAPFYTVPLQL 583
Db 596 DHTYAKQNALQVQVARGQPVVICDKEDTETIKTKRTIKVPHSVDCIQGILSVIPLQL 655
Qy 584 LAYHVALIKGTDVQDPRNLAKSVTVE 609
Db 656 LAFHLAVLRGYDVPFPRNLAKSVTVE 681

RESULT 10
US-10-480-021-8
; Sequence 8, Application US/10480021
; Publication No. US20060121457A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: GFATs AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-068
; CURRENT APPLICATION NUMBER: US/10/480, 021
; CURRENT FILING DATE: 2003-12-04
; PRIOR APPLICATION NUMBER: US 60/296, 076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328, 605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/357, 253
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 682
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-480-021-8

Query Match 35.8%; Score 1103.5; DB 6; Length 682;
Best Local Similarity 37.9%; Pred. No. 1.1e-73;
Matches 264; Conservative 118; Mismatches 213; Indels 101; Gaps 14;
```

```
Qy 1 MCGIVGAI-----AORDVAEILLEGRLRYRGYDSAGLAVVD-----AEGHMTFLRRL 49
Db 1 MCGIFAYMNYRVPRTKREIFETLIKGLQRLRYRGYDSAGVADIGNHNHEVKERHIIQLVKKR 60
```

```
Qy 50 GKVQMAOAA-----EEHPLHGGTGIATRWATHGEPSEVNAHPHSE---HI 94
Db 61 GKVKALDEELYKQDSMDLKVEFETH-----FGIATRWATHGVPASVNSHPORSKGNFF 115
Qy 95 VVHNGIENHEPLREELKARGYTFVSETDTEVIAHLVNMEL-----KQGGTLREAVLRAI 150
Db 116 VVHNGIITNYKDLKFKLESKYGEFESSETDTEIAKLKIVYFDNRETEDITPSTLVERVI 175
Qy 151 POLRGAYGTVMDSRHPDPTLLAARSGSPVLIGL-----
Db 176 QOLEGAFALVPKSVHVPGEAVATRRGSPLLIGVRSKYLSTEQIPILYRTCTLENVKNIC 235
Qy 184 -----GMG-----ENFIASQOLALLPVRTRFFIEEGDIAEITRRSVNIPDK 225
Db 236 KTRMKRLDSSACLHAGVDKAVEFFPASDASAVIEHTNRVIFLEDDDDIAAVADGKLSIH-- 293
Qy 226 TGAEYKR-----QDIESNLOVDAGDKGIYRHYMQKEIYEQPNAIKNTLTGRISHG 275
Db 294 --RVKRGASDDPSRAIQTLQMELOQIM--KGNFSAPMQKEIFEQPESVFNTMRGVNPF- 347
Qy 276 QVDLSSELGNDEL--LSKVEHIQILACGTSYNSGMVSRVWFSLAGIPCDVIEASEFRY 333
Db 348 EFTNTVLLGLKDHLEIRRCRLIVIGCGTSYHAAVATROVLEELTELPVMVELASDFLD 407
Qy 334 RKSAVRRNSLMTTISQSGEATDTLAGRLSKELGYLGLSLAICNVPGSSLVRESDLALMTN 393
Db 408 RNTVPFRDDVCEFTISQSGEATDTLLALRYCKRGAL--TVGTTNTVSGSISRETDGCVHIN 466
Qy 394 AGTEIGVASTKAFPTQLTVLLMLVAKLSRLKGLDASIEHDIVHGLQALPSRIEQMLSQDK 453
Db 467 AGPEIGVASTKAYTSQFISLVNFGMLMSSEDRISLQNRROEIIIRGLRSLPELIKEVLSLE 526
Qy 454 RIEALAEDFSDKHIALPLRGDDQVPIALEGALKLKEISYTHAEAYAGELKHGPLALIDA 513
Db 527 KIHDALELYTQSRLLVNGRGYNTATCLEGALKIKEITYMHSEGLAGELKHGPLALIDK 586
Qy 514 DMPVTIWAAPNNELLEKLKSNEEVRARGGQLYVPADQDAGFVSSDNHIIEMPHVEEVI 573
Db 587 QMPVTIMVKDPCFAKQNALQVQVARGQPIILCSKDDTSSKPAYKTIELPHVTVDCLQ 646
Qy 574 PIFTYVPLQLLAYHVALIKGTDVQDPRNLAKSVTVE 609
Db 647 GILSVIPLQLLSFHLAVLRGYDVPFPRNLAKSVTVE 682

RESULT 11
US-10-953-349-34174
; Sequence 34174, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: ENCODED THERBY
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 34174
; LENGTH: 680
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-34174

Query Match 35.3%; Score 1089.5; DB 6; Length 680;
Best Local Similarity 36.9%; Pred. No. 1.2e-72;
Matches 253; Conservative 123; Mismatches 228; Indels 81; Gaps 15;
```

```
Qy 1 MCGIVGAI-----AORDVAEILLEGRLRYRGYDSAGLAVVDAGEGHMT- 45
Db 1 MCGIFAYMNYRVSRRRYILEVFNGLRRLRYRGYDSAGIA-LDARDQVPSPPASSSDA 59
Qy 46 -----LRRLGKVMQLAAAEH-----PLHGGTGIATRWATHGEPSE 83
```


Db 1 MEVNRQLEGAYALFKSPHYPNELIACKRGSQLILGNVLSGQNGKSFHDVKTLTTNGK 60
QY 185 MGENFTASDQALLPVTRRFIFLEGDIAETRVSNI--FD---KTGAEVKRODIESNL 239
Db 61 PKELFFSSDLCAIVHTKNTYLALEDNEIVHIKDGVSILKFPDHPKPKPASVQRALSVMLE 120
QY 240 QYDAGDKGIYRHYWQKEIYEQPNAIKNTLTGRISHGOVDLSLGPNADELLSKVHEHIQIL 299
Db 121 EVGQIKGSDYHFMQKEIHEQPHSLKTTMRGRUKDGGVVLGGLKEYL-KTIRRCRRVFI 179
QY 300 ACCTS YNSGMVSRWFESLAGIPCDVEIASEFPRYKSAVRNLSMITLSQSGETADTLAG 359
Db 180 GCGTSYNAALAARPFVEELTGIPTVMEVASDLLDRQGPYREDTAVFVSQSGETADTLA 239
QY 360 LRLSKELGYLSLAIACNPVGSSILVRESDLALMTNAGTEIGVASTKFTQTLTVLLMLVAK 419
Db 240 LDYALENGAL-CVGITNTVSTLSRKTGCHGVHINAGCEIGVASTKAYTSQIVAMAMMALA 298
QY 420 LSLRLKGLDASIEHDIHVHGLQALPSRIEQMLSQDKRIEALAEFSDKHHLFLGRGDQYPI 479
Db 299 IGSQISTQARRDSIISGLNNLSSNVSEVLKLDAGMKELASSLIDSESLLVFGRGYNAT 358
QY 480 ALEGALKKEISYIHAEYAAGELKHGPLALIDADMPIVAVPNNNELLEKLSNIEBURA 539
Db 359 ALEGALKKVEALMHSEGMLAGEMKHGPLAVDENLPIIVIATRDACFSKQSQSVIQQLLS 418
QY 540 RGQLYVFADQ-DAGFV--SSDNMHIEMPHVEEVIAPFYTPVPLQLLAYHVALIKGTDV 596
Db 419 RRGRLVMSRGDASAVCPSSGGSCRVIEVPQVADCLQPVINIPLQLLAYHLTVLRGFDV 478
QY 597 DQPRNLAKSVTVE 609
Db 479 DQPRNLAKSVTTQ 491

RESULT 14
US-10-953-349-34176
; Sequence 34176, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; NUMBER OF SEQ ID NOS: 304
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 34176
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-34176

Query Match 26.2%; Score 807; DB 6; Length 488;
Best Local Similarity 37.1%; Pred. No. 6.1e-52;
Matches 181; Conservative 96; Mismatches 179; Indels 32; Gaps 7;

QY 152 QLRGAYGTIMDSRHPDTLAARSGSPVLVIGL-----MGENF 189
Db 3 QLEGAYALFKSPHYPNELIACKRGSQLILGNVLSGQNGKSFHDVKTLTTNGPKELF 62
QY 190 IASDQALLPVTRRFIFLEGDIAETRVSNI--FD---KTGAEVKRODIESNLQYDAG 244
Db 63 FSSDLCAIVHTKNTYLALEDNEIVHIKDGVSILKFPDHPKPKPASVQRALSVMLEVEQI 122
QY 245 DKGIVRHYWQKEIYEQPNAIKNTLTGRISHGOVDLSLGPNADELLSKVHEHIQILACCTS 304
Db 123 KGSYDHFHFMQKEIHEQPHSLKTTMRGRUKDGGVVLGGLKEYL-KTIRRCRRVFI 181
QY 305 YNSGMVSRWFESLAGIPCDVEIASEFPRYKSAVRNLSMITLSQSGETADTLAGLRLSK 364

Db 182 YNAALAARPFVEELTGIPTVMEVASDLLDRQGPYREDTAVFVSQSGETADTLALDYAL 241
QY 365 ELGYLGSLSLAICNPVGSSILVRESDLALMTNAGTEIGVASTKFTQTLTVLLMLVAKLSRLK 424
Db 242 ENGAL-CVGITNTVSTLSRKTGCHGVHINAGCEIGVASTKAYTSQIVAMAMMALAIGSDQ 300
QY 425 GUIDASIEHDIHVHGLQALPSRIEQMLSQDKRIEALAEFSDKHHLFLGRGDQYPIALEGA 484
Db 301 ISTQARRDSIISGLNNLSSNVSEVLKLDAGMKELASSLIDSESLLVFGRGYNATALEGA 360
QY 485 LKLKEISYIHAEYAAGELKHGPLALIDADMPIVAVPNNNELLEKLSNIEBVRAGGOL 544
Db 361 LKVKVEALMHSEGMLAGEMKHGPLAVDENLPIIVIATRDACFSKQSQSVIQQLLSRRGRL 420
QY 545 YVFADQ-DAGFV--SSDNMHIEMPHVEEVIAPFYTPVPLQLLAYHVALIKGTVDVDPEN 601
Db 421 IVMSRGDASAVCPSSGGSCRVIEVPQVADCLQPVINIPLQLLAYHLTVLRGFDVDPEN 480
QY 602 LAKSVTVE 609
Db 481 LAKSVTTQ 498

RESULT 15
US-10-488-015-17
; Sequence 17, Application US/10488015
; Publication No. US2006008902A1
; GENERAL INFORMATION:
; APPLICANT: STICHTING VOOR DE TECHNISCHE WETENSCHAPPEN
; APPLICANT: VAN DEN HONDEL, Cornelis
; APPLICANT: RAM, Arthur
; APPLICANT: DAMVELD, Robbert
; APPLICANT: ARENTSHORST, Mark
; TITLE OF INVENTION: Methods and Materials for the
; FILE REFERENCE: 13603PCTUS
; CURRENT APPLICATION NUMBER: US/10/488,015
; CURRENT FILING DATE: 2004-02-27
; PRIOR APPLICATION NUMBER: EP 01203423.7
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: PCT/EP02/09639
; PRIOR FILING DATE: 2002-08-28
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 464
; TYPE: PRT
; ORGANISM: Aspergillus niger
US-10-488-015-17

Query Match 19.1%; Score 589.5; DB 6; Length 464;
Best Local Similarity 34.1%; Pred. No. 7.5e-36;
Matches 157; Conservative 75; Mismatches 151; Indels 77; Gaps 9;

QY 96 VVHGIENHEPLREELKARGYTFVSETDTEVIAHLVNNELKQGGTLEAVL--RAIPOL 153
Db 9 VVHGIITNYKELKALLESKGFRETETDTECIAKLTXYLDQQDPDIDFTVLAKAVKEL 68
QY 154 RGAYGTIMDSRHPDTLAARSGSPVLVIGL----- 183
Db 69 EGAFGLLIKSVHYPHEVIAARKGSPLVIGVTRSRMKKVDFVDVSESDGPLPAEQASQNV 128
QY 184 -----GM-----GENFIASDQALLPVTRRFIFLEEG 210
Db 129 AIKKSATGLLAPPDKSLHRSQSRAFLSDGVPQPAEFLLSDPSAIVHTKTKVLYLEDD 188
QY 211 DIAETTRRSNIF-----DKTGAEVKRODIESNLQYDAGDKGIYRHYWQKEIYEQPNAIK 265
Db 189 DIAIHEQQLNIHRLTKDDGTNSVRAIQITIELEQEI--KGNFHFHFMQKEIFEQSVV 246
QY 266 NTLTGR--ISHGOVDLSLGPNADELLSKVBEHIQILACGTSYNSGMVSRWFESLAGIPC 323
Db 247 NTRGRILDVANKQVTLGGLRQYI-STIRRCRRRIIFVACGTSYHSCMAVRGVFEELTEIPI 305

| | | | |
|----|-----|---|-----|
| Qy | 324 | DVEIASEFRYRKSAVRNLSMITLSQSGETADTLAGRLSKELGYLGSLAICNVPSSLV | 383 |
| Db | 306 | SVELASDFLDQAPVFPEDDTCVFSQSGETADSLMALRYCLERGA | 364 |
| Qy | 384 | RESDLALMTNAGTEIGVASTKAF | 443 |
| Db | 365 | LLTHCGVHINAGPEIGVASTKAYTSQFVAMVFALSLEDRA | 424 |
| Qy | 444 | RIQMLSQDKRIEAL-AEDFSDKHAF | 482 |
| Db | 425 | QFKEILKLINEPIKOMCAKFFKNOKSLLLLGRGGQFPTALE | 464 |

Search completed: August 7, 2006, 09:31:32
Job time : 158 secs

This Page Blank (uspto)